



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 197923

TO: Ginny Portner
Location: REM/3B02/3C18
Art Unit: 1645
Monday, August 14, 2006
Case Serial Number: 10/780250

From: Vira David
Location: Biotech-Chem Library
REM-1A41
Phone: (571)272-1972

Virajita.David@uspto.gov

Search Notes

Examiner Portner,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Vira David
Intern
STIC Biotech/Chem Library
(571)272-1972

SEQID1

This Page Blank (uspto)

STIC-Biotech/ChemLib

197923

From: Portner, Ginny
Sent: Tuesday, August 08, 2006 12:47 PM
To: STIC-Biotech/ChemLib
Subject: 10/780,250

Please search SEQ ID No 1 (polypeptide amino acid sequence). thanks

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02; Mail Box 3C18
(571) 272-0862

8/8/2006

this Page Blank (uspto)

this Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2006, 20:21:12 ; Search time 44 Seconds
(without alignments)
1607.257 Million cell updates/sec

Title: US-10-780-250-1

Perfect score: 3770

Sequence: 1 EVKQENRLNSESSESSQGL.....TSTNGIKKILIRSKGYEIG 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR 80: *
2: PIR1: *
3: PIR2: *
4: PIR3: *
5: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3631	96.3	764	2	I39934	protective antigen
2	984	26.1	875	2	I40862	Iota toxin compone
3	235.5	6.2	192	2	I39933	cryptic protein -
4	235.5	6.2	204	2	G59104	hypothetical prote
5	209	5.5	4688	2	F82885	hypothetical prote
6	203	5.4	1227	2	C97033	uncharacterized pr
7	200	5.3	2529	2	B64635	toxin-like outer m
8	192	5.1	6713	2	B89921	hypothetical prote
9	188	5.0	4919	2	T31105	hypothetical prote
10	187	5.0	1302	1	JC6009	surface-located me
11	186.5	4.9	1072	2	A86827	hypothetical prote
12	183	4.9	4152	2	T31102	filamentous hemag
13	182	4.8	1837	2	T41023	probable nuclear p
14	180	4.8	752	2	G90599	hypothetical prote
15	179.5	4.8	2401	2	T28676	rhoptery protein -
16	179	4.7	2269	2	T28677	rhoptery protein -
17	177	4.7	1553	2	T18502	hypothetical prote
18	177	4.7	3724	2	T18427	hypothetical prote
19	176.5	4.7	2340	2	B71704	cell surface anti
20	175	4.6	1658	2	S55101	hypothetical prote
21	174	4.6	1033	2	T37715	actin-interacting
22	174	4.6	1516	2	E71619	RAD2 endonuclease
23	174	4.6	5005	2	F82884	hypothetical prote
24	172.5	4.6	1024	2	T30868	RhoA-binding prote
25	172	4.6	1939	2	T18372	repeat organellar
26	171.5	4.5	1802	2	S52611	TYB protein - Yeas
27	171.5	4.5	2399	2	H71879	toxin-like outer m
28	170.5	4.5	821	2	S67087	hypothetical prote
29	170.5	4.5	1365	2	T30822	Imp1 protein - Myc

30	170	4.5	2178	2	S55805	alpha-toxin - Clos
31	169.5	4.5	1051	2	T18351	Imp1 protein - Myc
32	169	4.5	1315	2	T28679	fibrinogen-binding
33	169	4.5	1465	2	S31262	TYB protein - Yeas
34	169	4.5	1635	2	A10452	hemolysin [Importe
35	169	4.5	1803	2	S56894	TYB protein - Yeas
36	168.5	4.5	979	2	J00894	P115 protein - Myc
37	168.5	4.5	1125	2	E90598	membrane nuclease,
38	168	4.5	769	2	F88870	serine proteinase
39	167.5	4.4	1218	2	U80069	hypothetical PI op
40	167.5	4.4	786	2	T18469	hypothetical prote
41	166.5	4.4	2523	2	T18477	hypothetical prote
42	166	4.4	1076	2	C70007	hypothetical prote
43	166	4.4	3316	2	C90538	hypothetical prote
44	165.5	4.4	2526	2	T20531	hypothetical prote
45	165	4.4	1308	2	E71622	probable membrane

ALIGNMENTS

RESULT 1
I39934
protective antigen precursor - Bacillus anthracis plasmid
C:Species: Bacillus anthracis
C:Date: 19-Jul-1996 #sequence revision 19-Jul-1996 #text_change 09-Jul-2004
C:Accession: I39934; S69160; F59104
R:Welkos, S.L.; Lowe, J.R.; Eden-McCuttchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.
Gene 69, 287-300, 1988
A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthrac
A:Reference number: I39933; PMID:89172073; PMID:3148491
A:Accession: I39934
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-764 <RES>
A:Cross-references: UNIPROT:P13423; UNIPARC:UPI00001E609; GB:M22589; NID:G143280; PIDN:f
R:Priddman, T.C.; Gordon, V.M.; Leppla, S.H.; Klimpel, K.R.; Birch, N.P.; Lob, Y.P.
Arch. Biochem. Biophys. 316, 5-13, 1995
A:Title: In vitro processing of anthrax toxin protective antigen by recombinant PC1 (SPC)
A:Reference number: S69160; PMID:95142670; PMID:7840657
A:Accession: S69160
A:Molecule type: protein
A:Residues: 197-202 <PRI>
A:Cross-references: UNIPARC:UPI0000179B0E
R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler,
J. Bacteriol. 181, 6509-6515, 1999
A:Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbori
A:Reference number: A59091; PMID:99445483; PMID:10515943
A:Accession: F59104
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-313, 'Q', 315-764 <OKI>
A:Cross-references: UNIPARC:UPI0000001332; GB:AF065404; NID:G4694216; PIDN:AAD32414.1; P
A:Experimental source: strain Sterne
A>Note: similar to anthrax toxin moiety, protective antigen, pagA formerly pag, plasmid I
C:Genetics:
A:Gene: pXOI-110
A:Genome: plasmid
C:Function:
Y active components edema factor or lethal factor; the complex is internalized by recept
C:Keywords: exotoxin
F:1-29/Domain: signal sequence #status predicted <Sig>
F:30-196/Domain: propeptide #status predicted <Pro>
F:197-202/Product: protective antigen #status experimental <Mat>

Query Match 96.3%; Score 3631; DB 2; Length 764;
Best Local Similarity 97.0%; Pred. No. 1.5e-174;
Matches 716; Conservative 5; Mismatches 11; Indels 6; Gaps 3;

DB 1 EVKQENRLNSESSESSQGLGYFSDLNFOAMVVTSTTGDLSPSSLENIPISENQYF 60
|||||
30 EVKQENRLNSESSESSQGLGYFSDLNFOAMVVTSTTGDLSPSSLENIPISENQYF 89

QY 61 OSATWSGFIKYKKSDREYTFATSDAHNTMTWVDDQEVINKASNSKIRLEKGRLOYKIOY 120
DB 90 QSATWSGFIKYKKSDREYTFATSDAHNTMTWVDDQEVINKASNSKIRLEKGRLOYKIOY 149
QY 121 QRENTEFGDLFKLYWTSQNKKEVISSDNLQLPELAKOSNSRKKRSTSGAPTYPDBDN 180
DB 150 QRENTEFGDLFKLYWTSQNKKEVISSDNLQLPELAKOSNSRKKRSTSGAPTYPDBDN 209
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTYKXSPEKWSGASDPYSDFEKYT 240
DB 210 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTYKXSPEKWSGASDPYSDFEKYT 269
QY 241 GRIDKNVSPPEARHPVLAAYPIYVHVMENIILSKNDQSTQNTDSTQRTISKRTSRRDAN 300
DB 270 GRIDKNVSPPEARHPVLAAYPIYVHVMENIILSKNDQSTQNTDSTQRTISKRTSRRDAN 328
QY 301 TVGVASISAGYONGF---TGNITTSAGFSNSNSTVAIDHSLSLAGERTMAETMGANTADT 357
DB 329 TSEVHGNAEVAHASFDPDIGSV--SAGFSNSNSTVAIDHSLSLAGERTMAETMGANTADT 386
QY 358 ARLANNIRYVNTGTAPINNVLETTSLVIGKQTLATIRAKENQLSQIILAPNNYPSKNLA 417
DB 387 ARLANNIRYVNTGTAPINNVLETTSLVIGKQTLATIRAKENQLSQIILAPNNYPSKNLA 446
QY 418 PIALNAODPSSPTITMNNYNOFLELEKTKQLRLTDQVYGNATATNFNGRRVDTGSNW 477
DB 447 PIALNAODPSSPTITMNNYNOFLELEKTKQLRLTDQVYGNATATNFNGRRVDTGSNW 506
QY 478 SEVLPOIDETTRAILFNGKDLNLYVERRIYAANVSPDETTKPDMLTKEALKIAFGNEPN 537
DB 507 SEVLPOIDETTRAILFNGKDLNLYVERRIYAANVSPDETTKPDMLTKEALKIAFGNEPN 566
QY 538 GNLQYQKODITEFDNPDQOTSQNIKNQLAELANATNIYTVLDKIKLANKNMILIRDKRPH 597
DB 567 GNLQYQKODITEFDNPDQOTSQNIKNQLAELANATNIYTVLDKIKLANKNMILIRDKRPH 626
QY 598 YDRNNIAVGADESUYKEAHREVINSTGBLLNIDKIRKILISGYIVAEDETEGKEVIN 657
DB 627 YDRNNIAVGADESUYKEAHREVINSTGBLLNIDKIRKILISGYIVAEDETEGKEVIN 686
QY 658 DRYDMLNLSLRQDCKTFIDPKCKYNDKPLVYISNPNKYVAVAKENTIIINPSENGDTS 717
DB 687 DRYDMLNLSLRQDCKTFIDPKCKYNDKPLVYISNPNKYVAVAKENTIIINPSENGDTS 746
QY 718 TNGIKKILIFSKKGYEIG 735
DB 747 TNGIKKILIFSKKGYEIG 764
RESULT 2
140862
Iota toxin component Ib - Clostridium perfringens
C:Species: Clostridium perfringens
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C:Accession: I40862; S42774
R:Perelle, S.; Gilbert, M.; Boquet, P.; Popoff, M.R.
Infect. Immun. 61, 5147-5156, 1993
A:Title: Characterization of Clostridium perfringens iota-toxin genes and expression in
A:Reference number: I40861; MUID:94041637; PMID:8225592
A:Accession: I40862
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-875 <RES>
A:Cross-references: UNIPROT:Q46221; UNIPARC:UPI000008491; EMBL:X73562; NID:g929031; PID

QY 60 FQSAIWGFIKYKKSDREYTFATSDAHNTMTWVDDQEVINKASNSKIRLEKGRLOYKIOY 119
DB 95 IKSTIRWGRILIPSEDGEYIILSTDR-NDVLMQINAKGDIK---TLKVMKKGOAYNIRIE 150
QY 120 YOREN-----PTEKGLDFKLYWTSQNKKEVISSDNLQLPELAKOSNSRKKRSTSG 170
DB 151 IQDNKLSIDNLSVP-----KLYW-ELANGKTVIPEBNLFPPDYSKIDEND----- 195
QY 171 AGPTVP-----DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIH 211
DB 196 --PPIPNNNFEDVAFPSAWEDEDLDTDNNDI PDAYENKNGTII---KDSIAVKNDSFA 249
QY 212 EKKGLTYKXSPEKWSGASDPYSDFEKYTGRIIDKNVSPPEARHPVLAAYPIYVHVMENIIL 271
DB 250 E-QGKTKVYSYLSNTRAGDPYTDYQKASGSIDAIKLEARDPLVAAYPVVGKHEKILIT 308
QY 272 SKNDQSTQNTDSTQRTISKRTSRRDAN--DANTVGVASISAGYONGFTGNITTSAGFSNSNS 330
DB 309 STNHHAS---DQKTVSRATNSKTDANTVGVASISAGYONGFTGNITTSYSHTTDNST 364
QY 331 TVAIDHSLSLAGERTMAETMGANTADTARLANNIRYVNTGTAPINNVLETTSLVIGKQTL 390
DB 365 AVODSN---GE-SWNTGLSINKGESAYINANVRVNTGTAPIMYKVPPTNLVL-DEET 417
QY 391 LATTAKENQLSQIILAPNNYPSKNLAPIALNAODPSSPTITMNNYNOFLELEKTKQLRL 450
DB 418 LATTAKENQLSQIILAPNNYPSKNLAPIALNAODPSSPTITMNNYNOFLELEKTKQLRL 477
QY 451 DTDQVYGNATATNFNGRRVDTGSNMSSEVLPQIETTRAILFNGKDLNLYVERRIYAANV 510
DB 478 ETTVYSNGYGRKN-SQGOI-ITBGNMSNYSIQSDVSASIIID-TSGQTFERRVAAEQ 534
QY 511 SDPLETTKPDMLTKEALKIAFGNEPNENILOY-QKODITE-FDNPQOTSQNIKNQLA 567
DB 535 GNPEDKT-PEITIGEAIKKAFSAATK-NGELIYFNGIPIDESCVELIPEDNSTSEIKEDQLK 592
QY 568 ELANATNIYTVLDKIKLANKNMILIRDKRF--HYDR-NNIANVGADESUYKEAHREVINST 624
DB 593 YLDRKCKLYNV---KLEKGMNILLIKVPSYFTNPEYNNFP--ASMSNIDTKNODQLQSV 646
QY 625 EGL-----LNIIDKIRKILISGY-----IVEIETEGLEKVINDRYDMLNI 665
DB 647 NKLGETKIIIPMKIKKRYKRYVFSYKSDSTNSITVINKSKQKTDIVLPEKDYKPF 706
QY 666 S-----SLRQDCKTFIDPKCKYND--KLPLVYISNPNKYV-----N 697
DB 707 SYEFETGKQSSDIETILTSSGVIFLDNLSTELNSPTEILKEPEIKVPSQDEIILDAHNK 766
QY 698 VYAVTKENTIIINPSENGDTSNGI 721
DB 767 YYADIKGLDT-----NTGNTYIDGI 785
RESULT 3
119933
cryptic protein - Bacillus anthracis
C:Species: Bacillus anthracis
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C:Accession: I19933
R:Welkoe, S.L.; Lowe, J.R.; Eder-McCuehan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.
Gene 69, 287-300, 1986
A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthrac
A:Reference number: I19933; MUID:89172073; PMID:3148491
A:Accession: I19933
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-192 <RES>
A:Cross-references: UNIPROT:P13422; UNIPARC:UPI0000168608; GB:M22589; NID:G143280; PIDN:

QY 1 EYKQENRLINSESSOGLIGYFSDLNQAPMVVYTSSTGDLSPSSLENIIPSE-NQY 59
DB 35 DTNQKEEITENTLSSNGLMGYIFADBEHFKDLEMAPIRKNGDLKFBKKVVDKLLTENDSS 94

Query Match 6.2%; Score 235.5; DB 2; Length 192;
Best Local Similarity 34.6%; Pred. No. 3e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

A:Molecule type: DNA
A:Residues: 1-1227 <KIR>
A:Cross-reference: UNIPROT:O97K41, UNIPARC:UPI00000CA0A1, GB:AE001437, PIDN:AAK79054.1,
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1080

Query Match 5.4%; Score 203; DB 2; Length 1227;
Best Local Similarity 19.5%; Pred. No. 0.02; Mismatches 285; Indels 280; Gaps 44;
Matches 170; Conservative 139;

```
37 STTGDLSPSELENIPISENQYFOSAIW-SGFIKYKSDDEYTPATSDNHTMWDQOE 95
14 SATTSAVALSKRAIAYAADNSVSSISNSNEIIVKGB-----VQEBE 58
96 V-INKASNSKIRLEKGRLYQIKIYOENPTK-----GLDFKLYWTDSONKKEVI 146
59 VVNNKNSNS-XYSSNSNQVS--NKNSNPKYSSSSEIQSINKNVNLYQVQNNKSVL 114
147 SSDNLQELBELKQKSSNSRK-----RSTSGPTVPDRNDGI PDSLSEVGYVDYKNNRT 201
115 AASNVDEVKLINSNVQTSYIAIGETKVKPDTLITNKALVDARSABG--TDLSEVEI 172
202 F-----LSPWISNI-----HEKKG-LTKYK 220
173 YDIYSQTAEAIQEAERINIDGVANVSDTLTGATFVNDANLDSVKKYFYHKRYATVTKF 232
221 SSPKWSYA-----SDPSDEKVTGRIDKRVSPKARPLVAA 258
233 DVATKTSNALKNINNGGERTDYALAEVSGVQPYLDL-----VNNKIIVKEQO----- 280
259 YPIVAVDMENIILSKNEPOS--TONPDSQRTSKSTSRDANTVGSISAGYQNGFTGN 317
281 -----NGRDLITTEISDSAASTIAR--INRALDMMDGAVATLEBYO----- 320
318 ITTSAGFSNSNSSTVAIHSLSLAGERTAEFM-----GLNTADTYRLNANI-----RYV 367
321 ---AIGANNVPLAHVADVNSLAM--DQRMGVSBADIGINTIMTYINNINSVGTEDEYI 375
368 NTGTAPI-----YVLEPTTSL--VLGKQNTLATIKAKENQLSQILAPNNY----- 411
376 NSHAVDSNEGNDIYDILNANIIEKTKAGQDL--TIPEVANVVEKVTLLDFYNHAAAGT 434
412 ---PSKNIAPIALNAQDDFSPTIMNYPNPLEKTKQLRLDPOV----- 455
435 TLQDYKAVDPNAQVQDDVAT-----LSDMLKTFDCKTKALQDKIDSILNSLKNINS 488
456 YGNIAITYNFENGRVAVDTGSNWSVYLPOIQETTARI--IFNGKOLMIVERRIAANVSPDL 514
489 IGNIDYS-----KLQTEAVDASKLBAVNDIDIKIKADKGRDLTLOELRDSVKKTIDYI 542
515 ETT-----KPDMLKALKIAF--GFNEPNCNLOYQKODITEPFPNDQOTSQNIK----- 563
543 NSTSNVSGKDGSDSYITIGIDGVEI--NIEFVERIKESGITI--TIEIKKVIIBRI 597
564 NQIARLAI--ATNITYVD-----KIKLNAKNMILLRDKKFEHVR--N 601
588 VQLSEVYRIYTVGTVDYKTLGINNVANDNNIYINMELKKKKQVYKLODIOTRVDNTIN 657
602 NI-----AVGADESV-----VKEAHEVINSSTEGLLNIDKIRKILSGYIVEIDT 649
658 NIDVINKIGAGAVAVSDYFNIGITDYODILDYVANADLKIQNYKOVDDIIEKVEAKISY 717
650 EGLK-----EVNDRYDMNLNSSLKODGKTIFDKKNDKLEPLY-----ISNRYK----- 695
718 BALMRINIGEAVTDDPKALGLTDI-----NDGLLLYATYTDLQNNKRYTADEV 764
696 -----VNVY--AVTKENTIINPSENGDTSTNGI 721
765 IARVQAQIEIYRALMQIN--LGAATTADYNTLGI 796
```

RESULT 7

B64635
toxin-like outer membrane protein HP0922 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: B64635
R:Tomb, J.F.; White, O.; Kervavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khajak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney, L. Nature 388, 539-547, 1997
A:Authors: Wallin, B.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: B64635
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2529 <TOM>
A:Cross-reference: UNIPROT:O25579, UNIPARC:UPI00000D3038, GB:AE000602, GB:AE000511, NID:

Query Match 5.3%; Score 200; DB 2; Length 2529;
Best Local Similarity 18.3%; Pred. No. 0.082;
Matches 172; Conservative 121; Mismatches 339; Indels 306; Gaps 38;

```
41 GDLSPSELENIPI--SENQYFOSAIWSGFIKYKSDDEYTPATSDNHTMWDQOEVI 97
508 GDLPL-KLKQLANSPTGFSEQAFNQAO-----KQEQDEQTLQ-----NEKTF 549
98 NKASNSKIRLEKGRLYQIKIYOENPTKGLDPLKYWTDSONKKEVISDNLQPLK 157
550 NSBOBGLKQAIQQAQAOQK--QQAQKQEQQAQYTYQEBLTHSQSLNVAADN-----TIA 604
158 QKSNSRKGRSTSGAPTPDRNDGI PDSLSEVGYVDV--KNNKTFSLPWISNHEKKG 216
605 SNDNTYNNQNTALIKEDAQGLNNTQQAQEQALQGLDLKLQQLANSF--TGFSEQAFN 662
217 TKYSSPEKMSYADPSYDEKVTGRIDK--NVSPEARHLVAAVPIVAVDMENIILSK 273
663 QAQKQEQDEQTLQNEKTFNSBOERLKQALANAKPTSPTPSHATPTPKHAPNTPKV 722
274 NEDSTONTDS-----QTRTISK-----NTSTSRDANTVGS 305
723 PPTPTQNPRAESVSGVYMLQNKITYSKGIYIDPNLSGQSGSGNTLSTYTNLFGRS 782
306 ISAGYONG-----FTGNITTSAGFSNSN 328
783 FSVNIQNTLLIIGNNTBSVNSGLIWHGGPGYITGFSAANIYLTNNPFTGEGVNSD 842
329 S-----STVAIDHSL-----SLAGRT 345
843 GGGANITFEKASDNITMDGLYNDAETVTKMLOTGASQSHSVATFPALNNISVTNSSFSDMT 902
346 W-----AETMGLMTAD-----TARLANAIRVNT--GTAPIYVLEPTTSLV 384
903 WKSFSFSKAKNISFSVNASGSGFTNPGSSVISAANATNSLSFINSRLNGAVNR--LQANSLI 961
385 LGKQNTLATI-----KAKENQLSQILAPNNYPS-----KNLAPIAL--NAQ 424
962 FNNQAVFNVLVYSGTSMFNATFTQLGNTNFTLSQSLLNFNGDTTLQNNANITLGNSQ 1021
425 DDFSTPTIMYNOFLEKTKQLRLDQYGNIAITYNFENGRVAVDTGSNWSVYLPOI 484
1022 AAFKNS--LTLNNSNSTLSDNOSVLANANTSAFNNQASLNTINGS----- 1064
485 QETTARIIFNGKDLNL--VERRIAANVPSDPLETTKPDMLKALKIAGFNPNGNLQY 543
1065 QATNSLFFNGGTLSLANASKLANSNASPSNNTT---INLDSVLSASNTSILANINIFQ 1121
544 GKDITF-----DENPDQOTSQNIKQLA----- 567
1122 GASQADFGGNTTIDTASFNPDSASLNFNNLTANGALNFNGYTPSLTKALMSVSGQFVIG 1181
568 ---ELNATNITYVDKIKLANAKNNILIDKRF-----HYDRN 601
1182 NNGDINISDI--NIFDNITKSVTYNLTNAQKITGISGANGYEKILFYGMKIQNAITYSDNN 1240
```


QY 602 NTAAGA-----DESVVKEAHR-----EYIN--SSTEGILNIDDI-----RKILSG 641
DB 1241 NQJWSEFNPINLSSQIIOESIKNGDLTIEVLNPNNSASNTIFNIAELNYYQASKNPFNG 1300
QY 642 YIYIEIDTGLKEVINDRDMNLISL-----RQDGEFIDFKYNDKPLY---IS 690
DB 1301 YKSDYSDNQA-----GTYLITSNIKGLFPYKSGTQAPCTTSPFNQPLASLNTYKGS 1355
QY 691 NPYKVVNVAVTKENTIIN---PSENGDSTNGIKIL 725
DB 1356 SENLKTLLGLHSQNSATLKEMIESNQLDNTN-INEVL 1392

RESULT 8

B89921
hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: B89921
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramoto, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; WUID:21311952; PMID:11418146
A/Accession: B89921
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-6713 <KUR>
A/Cross-references: UNIPROT:Q99U54; UNIPARC:UP1000011021A; GB:BA000018; PID:913701232; F
A/Experimental source: strain N315
C/Genetics:
A/Gene: ebha

Query Match 5.1%; Score 192; DB 2; Length 6713;
Best Local Similarity 20.3%; Pred. No. 0.87;
Matches 170; Conservative 128; Mismatches 297; Indels 244; Gaps 42;
QY 9 LNESSSSQGLLGYPYSDINFPAPVVTSTTTGDLSPSELENIPESENQYFQSA 63
DB 2468 VQQAASDAKANIIG-TLTHLNNAQKODLTQIBGATTVNGNSVKTAKODLDGAMQRLBSA 2526
QY 64 IWSGFIKVKSDY-----TFATSDNHYTMWVDDQEVINKASNSKIRLEKRLYOIKI 118
DB 2527 I-ANKDQTKASENYIDADPTKTAFFNALT---QAESYLNKHGHTKMDK-----QAVQ 2576
QY 119 QYQRENPEKGL--DKLYWTSQNKKEVYSSNQLDPE---LKQSSNSRKRKSTASAP 173
DB 2577 AIQSVTSTENALNGDNLCCAKTEATGAILDNLTLPTPQTKALKQVNAQR----- 2628
QY 174 TVPDRNDGIPDSLEVEGYTVDKNRTFLSPWISNIHEK-----KGLTKYKSSPEK 225
DB 2629 -----VSGT-DLKSATSLNANMDQLKALIGHDITVAGANTASPRK 2672
QY 226 WSTASDPY-----SDFEKVTGRI-----DKVSPBAHPVAAV 259
DB 2673 QGAYTAYNAAKNIWNGSPVITNADVTAATQRVNNAETSLNGDTNLATAKQADALR 2732
QY 260 PIVHV-DMENIILSKNEDSTQNTDSQTRISKSTSRD-----ANTVGSISA 308
DB 2733 QMTHLSDAQKQSTQIDSAITQVTVQVOS---VRDNATNLDNANMOLRNSIANDDEVKASQ 2789
QY 309 GY-----QNGFTGNTSAGFSNSSTVALDHLISLAGERTWAEWGLN----- 353
DB 2790 PLYDADPTDQNAVTAVTSAGENIINATSOPTLDPASVTAQAAQVNTNKTALNKAQVLANK 2849
QY 354 ----TADTARL-----NANIRYVNTGAPLIYVLPPTSLVGNQTLATIKAKENQSLI 405
DB 2850 KQETANINRLSHLNNAAQKODLNTQVTAAPNI-----STVQVYTKAEQLDQAM 2898
QY 406 APNNYTPSKRLAIALNADDFSTPTTNNVYQFLSEKTKQALRLTDQYVGNIAF----- 461

DB 2899 ER-----LINGIQDKQVQSVN---FTDADPEKQ-----TAYNNAVTALEN 2937
QY 462 -YNEFNGRVAVDTGSMSEVLPOIQE-TTARILFNKDLNLYERRI--AAVNSDLEF- 516
DB 2938 IINQANG-----TNANQSOVEBALSTVTTTKQALNG-----DRKYTDKNNANQTLSTL 2986
QY 517 -----TKPDWTLKEALKTAFGEFENGNIQ--YQKDIPEPDFNF 555
DB 2987 DNLNNAQKAVTGNINQAHVAVT--QALQTAQELNTAMGILKSLNDKDTLLSGQNTA 3044
QY 556 -----OOTSQNIKNQALNATNITYVLDKIKNAKXNILLRDRPHYDRNNIAV 605
DB 3045 DADPEKKAAYNAVRNAENILNKSSTGTN-----PKQVBAAN-----QVNTTKAL 3092
QY 606 GADESIVKEAHRREVINSSTEG--LNLNDKIRKILSGTIVEIEDTEGLKEVINDRYDL 663
DB 3093 NGTQNL--EKAQKHANTALDGLSHLTNAQKALKQLVQOSTVVAEQNGEOKAN----- 3144
QY 664 NISS-----LRO--DQKTFIDFKYNDKPLYISNPNYKVVNY--AVTKENTIPSEN 713
DB 3145 NVDAMDKLRQSIADNATTKQONQNTD-----ASFN-KKDAVNNAVTTAQGIIDQTTN 3196

RESULT 9

T31105
hypothetical protein 2 - Haemophilus ducreyi
C/Species: Haemophilus ducreyi
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C/Accession: T31105
R/Ward, C.K.; Lumley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.
J. Bacteriol. 180, 6013-6022, 1998
A/Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.
A/Reference number: Z20984; WUID:95030326; PMID:9611662
A/Accession: T31105
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-4919 <KAB>
A/Cross-references: UNIPROT:Q9ZHL0; UNIPARC:UP1000011020B; EMBL:AF057696; NID:93929021; F
A/Genetics:
A/Gene: lspa2

Query Match 5.0%; Score 188; DB 2; Length 4919;
Best Local Similarity 20.5%; Pred. No. 0.87;
Matches 187; Conservative 125; Mismatches 323; Indels 276; Gaps 43;
QY 6 NLLNESSSSQGLLGYPYSDINFPAPVVTSTTTGDL-SIPSELENIPESENQYF---Q 61
DB 154 NNSIKKGNQVLGLGEN--KNIGSQAAKTIPIQVNTDQSKISGGLVEGKADFLIYN 212
QY 62 SAWSGFIKVKSDYTFATST----- 82
DB 213 NGVTLNGVKTINTDRAVASTSEVEPHIKQLNVORGVILIGQVATNGLSHEDVAAKNI 272
QY 83 -----ADNHYTMWVDDQEVINKASNSKIRLEKRLYOIKI-- 118
DB 273 EQQKVSIEGSKPAKLANVTAPAGNLTYDVNTRD--NNNTNPKKPIITDNTKONIAISG 331
QY 119 -----QYQREN--TEKGL-----DFKLYWTSQNK-----KEV 145
DB 332 EAGASGYGNIFITYLTKAGVNHQGVIFAEDDINILTDGNSRLKLVYADYRVVVGKDI 391
QY 146 ISSDNLQELKQKSSNSRKRKSTAGPTVPDRNDGIPD-SLEVEGYTVVKNKRTFLS 204
DB 392 ELANNQIHLADQOLILNATGHVKNLNDGSSVINSNNIGISALMLTLENATVSAAN---LS 447
QY 205 PWISNIHEKKGKLTKYKSSPEKWTASDPYD-----PEKYTGRI-DKQVSPBAHP 254
DB 448 FRYVNTDKLNLSKVS-----ARAADLOSGLNLDKASVLAHKLTLNISNDVSLANQSK 501
QY 255 LVAAVPIVHVDENIILSKNEDSTQNTDSQTRISKSTSRDANTVGSISAGYQNGF 314
DB 502 LSA-----NNLKIKVRDLNLSNSELNANNTLNTSN-----NITLKNKSKF 543

QY 315 T-GNITTSAGFSNSNSTVAIDHS:SLAGEPTWAEIMGANTADTARLANIRYVGTAP 373
DB 544 TAGNNTLVN---TNNVLANDSELA-ANNITLVNTKAVTLNDASKLSANKLIDLVN---- 594
QY 374 IYNNVPTTSVLVGNQTLA----TIKAKEN---QLSQTILAPNNYPSKMLAPIALNAODD 426
DB 595 -----TNNVLNNSKSTLSAGELTFKVKYKNTVLNNDSELAANN-----LSLNA--- 636
QY 427 FSSPTPTMANYNOFLEKLT--KQLRL---DTDYQYGNIAATTNFENGVRVVDTSNMSEVL 481
DB 637 --SHNVTLNNSKLSAQKADIKAVNLTLNDTTE---LTAKLIDINSTYTITNGTIGIF 690
QY 482 PQIOETTARIIIFNGKOLNVERRI--AAVNPSPLETTPEMDTLKEALKIAPGPNPNGL 540
DB 691 ANI---TTEKLNKKEKALILABQNLNFTVNGSH--YENKDIVSKDKATVTSKNS----- 741
QY 541 QYOGKDITEPFPNPDQOTSQNIKNQLAELNATNI-----YTVLDKIKLNK----- 586
DB 742 -----DTSNGSKLVNQNOL-KYNNVNFITISQDDTTLIGNVTLNNSGFTNSG 790
QY 587 -----MNLIRDKRPHYDRNNIYAGAD--BSVVK-----EAREVINS- 622
DB 791 NLTVTKTLVDGDIQWFTKKNLTVGEDLHKSKTKITNDGKLISIKNLNISSEADPFING 850
QY 623 -----STEGLLNIDKDIRKILSGYIYEIDTGLKEVINDRYDM--LNISS-- 667
DB 861 TLLGLEALKIKATKGFPTNEKAI--LASNSLIDISVASKGTFENGTJBSGNLITWNG 908
QY 668 --LRDQGTPIIDFKKYNDKLPYISN-----PNYKVNVAVTKENTIIINPSENGDSTNG 720
DB 909 AFLANDNNTTISFGVNLNTSTGNSVNNGTJLSNEBLN---ITSANFTNES-NGTVMGNG 964
QY 721 IKKILIFSKKG 731
DB 965 LINTI--AKOG 973

RESULT 10
JC6009
surface-located membrane protein Imp3 precursor - Mycoplasma hominis
C/Species: Mycoplasma hominis
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: J06009
R/Labeloged: S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christensen, G.
J. Bacteriol. 178, 2775-2784, 1996
A/Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis Imp gene system
A/Reference number: JC6009; MUID:96213016; PMID:8631664
A/Accession: JC6009
A/Molecule type: DNA
A/Residues: 1-1302 <LAD>
A/Cross-references: UNIPROT:Q49547; UNIPARC:UPI00000B0501; EMBL:X95601; NID:g1197335; PI
C/Genetics:
A/Gene: Imp3
A/Genetic code: SGC3
C/Superfamily: surface-located membrane protein Imp3; tetratricopeptide repeat homology
C/Keywords: duplication; membrane protein
F/1-24/Domain: signal sequence #status predicted <STG>
F/25-1302/Product: surface-located membrane protein Imp3 #status predicted <MAT>
F/957-992/Domain: tetratricopeptide repeat homology <TT1>
F/993-1026/Domain: tetratricopeptide repeat homology <TT2>
F/1089-1120/Domain: tetratricopeptide repeat homology <TT3>
F/1154-1190/Domain: tetratricopeptide repeat homology <TT4>

Query Match 5.0%; Score 187; DB 1; Length 1302;
Best Local Similarity 18.7%; Pred. No. 0.14;
Matches 141; Conservative 139; Mismatches 296; Indels 178; Gaps 28;
QY 1 EYKQENRLNBESSSQGLGYFSDLNQAPMVYTSSTGSLSPSSLENIPEBNQYF 60
DB 476 QYDEANKSIKEQJNLIDKANTLLPOLNDNSEIYAKESILNAETTNAKAVNQNDNSM 535
QY 61 OSAINSGFIKVKKSDEYTFATSDNHVTMWMVDQEVINKASNSKILREKGR--LYQIKI 118

DB 536 QSA-----KSLDDKVTYKIQNLTEBFNKDKAKFELEQTRKIDNFLT 579
QY 119 QYQRENPEKGL-----DFKIYWDSONKKEVISDNIQLBELQKSSNR--KGRS 168
DB 580 DDVKNNPVATLVKDLTNAKDDKSVTKSSNKSIIAAND---ELKQALDKARVAQOI 635
QY 169 TSAGPTVDRNDGIPDSLEVEGYTV---DVKNKRFTLSPWISNIHEKGL---TKYKS 221
DB 636 DEANKSIKEQJSDSITTNANQNLNLVDSKDIOQAKTELSQSEIGASQSEELNPNTSQMS 695
QY 222 SPE---KMSASDPYSDPEFKVTRIDKNVSPPEARHPLVAAYPIVHWMENIILSKNEDQ 277
DB 696 AKESLDAKVTEITTKLETFNK-----DKDYK-----FKLEKTRKIDIEFI 736
QY 278 STQNTDSQTRISQKTSISRPANTVGVSISAGYONGFTGNITTSAGPSNSSTVAIDHS 337
DB 737 NTKNTNPVYSTLISELTSKRDSK-----NSVTNSNSNSDIETANTELKQA 781
QY 338 LSLAG-ERTWAEIMGANTADTARLANIRYVGTAPIYVLPFTSLVLGKNQTLATIKA 396
DB 782 LAKANTDKAQDNIAKSTKE--QLNNSISSANTLLAKLTD-----KDNITIQAKT 829
QY 397 ---KENQLSQTILAPNNYPSKMLAPIALNAQ-DDFSSTPTMANYNOFLEKTKQRLDT 452
DB 830 ELEKRVQKQAVASNNNTASQSAKSLDAKVTEITTKLETFNKDKDYKFELEQTRKDI 889
QY 453 DQYGNIAATTNFENGVRVVDTSNMSEVLPOIOETTARIIIFNGKOLNVERRIAAVNSD 512
DB 890 DEF-----INTNKNPDYSTLISELTSK-----RDSKNSTNNSN 924
QY 513 PLETTKPMTLKEALKIAFGNPNPNQYOGKDITEPFPDQOTSQNIKNQLAELNAT 572
DB 925 KSDIETANTEKQALARA-----NTBKQDQAD-NLASTGEQLNKSISSAN-- 968
QY 573 NIYTVLDKIKLNANKNMLIRDKRPHYDRNNIYAGAD--BSVKEAREVINSSTEGILL 629
DB 969 ---TLAKLT-----DKDNITIQAKTELEKRVQKQAVASNNNTASW-- 1007
QY 630 NIDKIRKILSGYIYEIDTGLKEVINDR-YDMLNLSLQDGTPIIDFKKYNDKLPY 688
DB 1008 ---OSAKSLDAKYTEI--TYKLETFNKDKVYKRELEQTRKIDIEFTNTNK----- 1054
QY 689 ISNPYKVNVAVT-----KENTIIINPSENGDST 718
DB 1055 -TNPYSTLISELTSKRDSKSNSTNSKSDIET 1087

RESULT 11
A86827
hypothetical protein yufg [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C/Accession: A86827
R/Bolotin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
A/Reference number: A86825; MUID:21235186; PMID:11337471
A/Molecule type: DNA
A/Residues: 1-1072 <STO>
A/Cross-references: UNIPROT:Q9CF64; UNIPARC:UPI00000D4491; GB:AE005176; PDB:g12724625; PI
A/Experimental source: strain IL1403
C/Genetics:
A/Gene: yufg

Query Match 4.9%; Score 186.5; DB 2; Length 1072;
Best Local Similarity 20.0%; Pred. No. 0.11;
Matches 163; Conservative 135; Mismatches 353; Indels 165; Gaps 30;
QY 2 VKQENRLNBESSSQGLGYFSDLNQAPMVYTSSTGSLSPSSLEN----- 52
DB 300 VEQYDVVASSESTQANSAKSLPISSEASVTNTNLSISLSDSISSSQTENSQSGASSTA 359

QY 53 ----IPENQYFQSA---IMSGFIKYSDBEYTPATSDAHNVYMWDDQEVINKASNSKI 106
 DB 360 EISYSDENSNLSLSSNQINSNSNSEKDSQSSSLGSSMSNESEHSSNSININETNNSSSEI 419
 QY 107 R-----LEKRLVQIKIYOORENPEKGLDFKL-----YWDSONKKEVISDMLQJPELK 157
 DB 420 TALLPSNPRESNVSDQTSSEASTNSNSISLSPSNISSTSEATNSDSNAEVA 479
 QY 158 QKSSNRKRKSTAGPTVPDRNDGI-----PDSLEVEGYTVDKRKFTLSPWISNIHEK 213
 DB 480 NNSLAVNNSSSSVLSTSTADNLGINQSGSDULTD--SSSISTSGAFLS---SNQTS 534
 QY 214 KGLTKKSPKMKSTADPYSDPEKYTGRIID-KVNSPEARHPVAAPIYVHMENIILS 272
 DB 535 EASYNSSSISLSPSNISSTSVLESTSSSPSNVAVANNISLASV-----580
 QY 273 KQEDOSTONTDQTRTISKY-----TSTRDANTVGVISAGYONGFTGNITTSAGFS 325
 DB 581 NMSSSSVLSTSTADNLGINQSGSDULTD--SSSISTSGAFLS-----NOTSSEASS 632
 QY 326 NSNSSTVAIDHSLSLAGERTMAETMGLNTADTARLAN-----IRYVNTGT-----371
 DB 633 NSMSSTNSPSSLISLSSNSESA-TNQSNSEATKVDNNSSTHSSNLTSSSDSDSD 691
 QY 372 --APIYVLEPTTSLVYKQOTLATITAKENQLSQ---IIAPNNYPSKNLAPIALNAQD 425
 DB 692 SPSDSNL--SSSPNLETNOTTSSKPSKPSVNNISENPKVSSNSVOENSTDHMSSTPKS 749
 QY 426 DPGSTPTTNNYNOFLEKTKQRLDTPQVYGNLATY--PENGVRVDTGSMSEVLP 482
 DB 750 SSSPSTSTSSSQKESQSN--LMTTSGINPITFNSSSENSASALT--SYSNNS 804
 QY 483 QIQTETARIIF-----NGKDL-----NLVERIAVANPSDPLETTEKDMTLK-EALK 528
 DB 805 ESEETCLVTSNKAQDNSEISHSLPSNSNENNVSSISQSAIILSSKSTTKRSSLS 864
 QY 529 IAFGNFEP-NGNIQYQKQDITEFDNPDQOSTONIKQALMALATNIIYVLDKI---KAN 564
 DB 865 IINSTSHPNQEDQNSNSD-----EYKSNNNVESIILQOLNSINKTKMNSLTQSKS 916
 QY 585 AKNNILIRKRFHYDNNNIAVGADSVVKAHREVNSTTEGLLNIIDQIRKILSGYIV 644
 DB 917 VYTLPSKSKVTEKNENSNVSEKLIKTPQKN-DESONLQGITALDLSFNK-----968
 QY 645 EIEDTEGLKEVIN-----DRYDMNLTSSLROD-GKTFIDPKKNDKLEIYISNPYK 695
 DB 969 EYETMEDSKTVPKULDNENGRSQNNKSTIAKDKNKKPFKRSERNSKIL-----1019
 QY 696 VNVYAVTKENTIIINPSNGDTSTNGIKLILFSGKG 731
 DB 1020 -----DSDNNILKKTIVLKKHG 1036

RESULT 12

T31102
 filamentous hemagglutinin 1 - Haemophilus ducreyi
 C:Species: Haemophilus ducreyi
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T31102
 R:Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.
 J. Bacteriol. 180, 6013-6022, 1998
 A>Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.
 A:Reference number: Z20984; MUID:99030326; PMID:9811662
 A:Accession: T31102
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-4152 <WAB>
 A:Cross-references: UNIPROT:Q9ZHU3; UNIPARC:UP100001101PD; EMBL:AF057695; NID:93929017;
 C:Genetic:
 A:Gene: LepA

Query Match 4.9%; Score 183; DB 2; Length 4152;

Best Local Similarity 21.4%; Pred. No. 1.2;
 Matches 177; Conservative 129; Mismatches 336; Indels 186; Gaps 43;
 QY 10 NSESSSSQGLGYFFDLNFPQ-APWVVTSTTGD-----LSPSSELENIIPENQYF 60
 DB 162 SRSTSTQ-LVGLHANIOLOQKAKILINQVGDHESNIQALAEVAGKADLIIVN-----216
 QY 61 QSAIWSGFIKYSDBEYTPATSDAHNVYMWDDQEVINKASN-----SKIRLEKRLV 114
 DB 217 PNGITLNGVKTINTDFEYVTS-----DIIIPRENGLSVRNGKVTITDKGVA 264
 QY 115 QIKIYQ---QRENPEKGLDFKLWYTDSONKKEVISD-----NILQPELKQSS 161
 DB 265 TGLSHFEVARNIIOQG--KITVAKTENQGSVNPANITPAAGSLNLIKTRPAPISS 321
 QY 162 NSRKRKSTAGPTVPDRNDGIIPDSLEVEGYTVD---YKRRFTLSPWISNIHEKGLT 217
 DB 322 GTSR---TSDTPAL-SADAGSWYGSNIKFWVTDKAGVYKHGIIIPSENDINIKMGQNA 377
 QY 218 KYKSPKRS-----TASDPISDEKY---TGRIDKNSPEARHPVAAPIYH 263
 DB 378 SLKEIYAKQDIDILANDIBLTERGQLOANNKIIINSTGKIN-----LNASEVSAADNVN 431
 QY 264 VDMENIILSKEDOSTONTDQTRTISKNTSRTDANTVGVSI-----SAGYONGFTG 316
 DB 432 VASENLAL-ENASMSANSILDVYTKIEVRSKVSAGTANIRASNTTLDGSSVVAANKIL 490
 QY 317 NITTSAGFNSN-----SSTVAIDHSLSLAGERTMAETMGLN-----TADTA 358
 DB 491 NYTNATNLNQSGLSAKQDELAVTNHITLNTSKLSAQANIKTEMLTLNGEASVLAERL 550
 QY 359 KLNANIRYVNTGAPRYVNLPTTSLVYKQOTLATITAKENQLSQIAPNNY-----YRS 413
 DB 551 DINAIQKITNGIAGLTANITTKAL--ENRDNALILAHQ--LNTFVNGSHYVNNKGDIVS 607
 QY 414 KNLAPIALNAQDPS-----TPITMNYNOFLEKTKQRLDTPQVYGNI---AT 461
 DB 608 KQDAIYTFESNSDFTSNGSKLYDAQNNLTVYVANNPFIITGSEI---LHENVTLANK 661
 QY 462 YNEFN-GRVRVDTGSMSEVLPQIQTETARIIFNGKDLNVER-----RIAAVNP 510
 DB 662 GNFTNSGNLTWKELNTSIESPI--NAGNLTGKGLVHSVTVYKNDGKLVSIENLNI 718
 QY 511 SPLEETKPDMLT-KEALKIAPG--FNEPENGLOQYQDITEFDNF--DQOTSQNIKQ 565
 DB 719 SSKTPTNNGTTLGLALAKIAGGAFNNAAGSLA-SNKSLLDYGNFTNGTIBSVKS- 776
 QY 566 LAELNATNIVYLDK--ILNNAKNNILIRKRFHYDRNNIIVAGAD-ESVYKAHREVINS 622
 DB 777 ---LNTITNNYTPINNAVTISYGVNLITSSQ-NFTNDSNGTVMSHDLNITTSQAN--TINK 830
 QY 623 ST-----EGLL-----NIDQIRKILSGYIVLEIEDTEGLKEVINDRYDMNLTSSL-RQDGK 673
 DB 831 NLLAGQGLNLTAKGNITMDSNTALAVLHNSNDIN-----LNANKYVNIIGETIYQAGN 885
 QY 674 TFIDFKYNDKLPYISNPNYKVVAVYKENTIIINPSNGDTSTNGI 721
 DB 886 ISVEAKLIHNDVLT-----SGNITVTK-----SGNAATVKTINSI 919

RESULT 13

T41023
 probable nuclear pore complex-associated protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T41023
 R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Bartell, B.G.
 submitted to the EMBL Data Library, June 1998
 A:Reference number: Z21965
 A:Accession: T41023
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1837 <MUR>

A:Cross-references: UNIPROT:O74424; UNIPARC:UP1000006A0P8; EMBL:AL023860; P1DN:CAA19588.

A:Experimental source: strain 972h-; cosmid cl62

C:Genetics:

A:Gene: SPDB:SPCC162.08c

A:Map position: 3

Query Match 4.8%; Score 182; DB 2; Length 1837;

Best Local Similarity 19.2%; Pred. No. 0.41;

Matches 168; Conservative 155; Mismatches 347; Indels 206; Gaps 39;

```
QY 1 EVKQENRLINSESSQGLGYFSDLNFGQAPVVTSTTGDSLSPSELENIPSENOYF 60
DB 773 QLEGNHNIYDKLSBNLLTST-----VKDKLKADLSMLEKSLSSIQDNEPHM 819
QY 61 QSAIWSGRTKXKSBDEYTPAT-SADNHTMWDDEVINK-----ASNSNKRLEKSG--- 111
DB 820 KAQIES-----SNQBYATYDSMNSRILELSNDLRVANSKLSSECSDDVRRLTLONSFDL 873
QY 112 RLVOIKIQORENPEPEKGLDPFLYWTDSQNKKEVISDNLOLPBLKOKSSNKRKRSTSA 171
DB 874 REHQTLVLVQLOSNITELKODITLQRT-VRNQLEI-----QTELEKRLKFMEEQENIQ 926
QY 172 GPTV-PDRNDGIPDLSLEVGYTVDV---KMKRTFLSPWISNIHEKGLTKYKSSPEKMS 227
DB 927 SKLIANKDQTQNPONVEVBALISIELBRTKEKLRMABLEKSNIQO-----KYLASEKTLR 981
QY 228 TASDPYSDF-----EKVTG-----RID-----KNSPEARHPLVAA- 258
DB 982 MNNEHTEQFKHLVESEISTREBEKITSLRSELDLNRKVEVLKEEKSESSKEIAKOLEDAV 1041
QY 259 -----YPIVHVMENIILSKNEB-----QSTQNTDSQRTISKNTSTSR 297
DB 1042 REKOSALSPKQDYKIRSDADRVITSLKEDIKERSLMKEGSHNSVESRIVSHGRTOQKLR 1101
QY 298 D-----ANTVGSISAGYQ---NGFTG---NITTSa-----GFSNSNS 329
DB 1102 DLRTFDEVTNTKYLKLNKFNBOQSGLSGAEKDMNIQKAKMEDISLSKDYILGLENOQNK 1161
QY 330 STVAIDHSL-SLAGERTMAETMGANTADTANLANIRYV-NTGAPPIYNYLPTTSVLV- 385
DB 1162 ----LHSPQPSLSQOIT---VLQONSSENLNISANLEAVQONDRLRELVSYLRIHEKEITMD 1214
QY 386 -----GKNQTLATIKAKENQLSQTILAPNNYVPSKNIAPIALNAQDPFSSTPTM 434
DB 1215 NKYEITIIDNGLNQVKSIGSTYDLSQLELNRLOSLSPVSN-----DQTTPTI-I 1263
QY 435 NYNQPLEL-EKTKQRLDPTDQVYGNIAITYFENGRVRVDGSSNNSSEVLPOIQETIARI 492
DB 1264 SSGSQVQALLYESNSVLRKONDAKLQELKEVEKEL---NASINPLQTEINELKAELIG 1319
QY 493 FNGKDLNIVE-----RRIAAVNSPDLPTTKYDMLTKEALKIAPGNEBNGNLQYOG 544
DB 1320 AKTASINLMKEVNSRWKLRFOGVLNKYERVDPTQ---LEELKNCCEALEKKOBELETKL 1375
QY 545 KDITFEDEFNDQOTS-----ONIKQOLAEANA-TNIYTVADKIKLANKNMILIRDRKF 596
DB 1376 QETAKETOTFKQVUNSLBEVENLKKVGEQANTKTRILAAANEKCEMLKSSSL---TRF 1432
QY 597 HYDRNNIIVGADSVVKEAREVINSSTEG-----LLNLIDKDIRKILSGYIVEI--- 646
DB 1433 AHLKQELTNKNKELTNSKNAENAMQKEISLKDSSHQOESASDAEQITKQFQOLKSE 1492
QY 647 -EDTSGLKEVINDRIDMLNIS--LRQDGKTFID-----FKYNDKLPYIYSNRYKYN 697
DB 1493 KERTS--KELADSKNELEHLQSEAVDADQKTEISMLEKEIHRLRSDKGLVOQVONLSAE 1550
QY 698 VYAVTKENTTIINSENGPTSTNGIKILIPSKGYE 733
DB 1551 LAALREHSPQOSLNADEIAR-LASQLESTQYIE 1585
```

hypothetical protein MYPU 7030 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C:Dates: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: G90599
R:Chamblaud, I.; Hellig, R.; Perris, S.; Barbe, V.; Samson, D.; Gallissou, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A9512; MUID:21267165; PMID:11353084
A:Accession: G90599
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-752 <KUR>
A:Cross-references: UNIPROT:Q98PL9; UNIPARC:UP100000C80F4; GB:AL445566; PID:g14090118; P1
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU 7030
A:Genetic code: SGC3

Query Match 4.8%; Score 180; DB 2; Length 752;

Best Local Similarity 19.2%; Pred. No. 0.14;

Matches 159; Conservative 130; Mismatches 268; Indels 272; Gaps 42;

```
QY 41 GDLSPSELENIPSENOYFQSAIWSGRTKXKSBDEYTF-----ATSADNHTMWD 92
DB 20 GTISGISYGIQLPNBSASL-----VKRANLEMFNLKNSYLNKSIIBELINQMOD 69
QY 93 DQEVNKAASN-----SNKRLEKGR--LYOI---KIQYORBNTEKGLDPFLKYWTDSQ 140
DB 70 ENKSNISANDPKNKVFQSKTPLENGEKITYSVAGKDIYFOIYNBSQTKISKSVKITSK 129
QY 141 NKEVISSDNLQPELKOKSSN--SRKRSTSGAGTVDPRND-----GIPDLSLEVGYTV 194
DB 130 ISKQVWMD-----KQRLNDFAKNLRVNFKSSASBEKSDIWAQFQNXSKLEIKYL 182
QY 195 DVKN-----KTFLSPMISNIHEK---GLTKYKSSPEKWSIASDPY 233
DB 183 DDKRVKKNISENTEBDEVELKTENNAFVGYGNDIAGTVIETAIYKYNSEK-----Y 235
QY 234 SDFEYVNRIDKN-----VSPEARHPLVAAPIYVHVMENIILSKNEQSTQNTDSQTRT 288
DB 236 QNIQKI-QITNNEFRPTSDS-----LMWELSNKVFEEQSKSEN- 276
QY 289 ISKNTSRDANTVGSISAGYONGFTGNITTSAGFSNSNSTVAIDHSLSLAGERTWAE 348
DB 277 -----DA-----SAFSGNSLVSKNSLKNQ-----DEKIKLBS--VF 307
QY 349 TNGIANTADTANLANIRYVNTGAPPIYNYLPTTSVLVGRNOTLATIKK-----ENQ 400
DB 308 VFDELKNTNENTKLSYKTRDIK-----FNKLETVBQKGSVRLTYLIGKVVYDGNESIPNR 361
QY 401 LSQILAPNNYVPSKNIAPIALNAQDPFSSTPIIMNYYNQFLELEKTKQRLDPTDQYGNIA 460
DB 362 LTLSPRPS---SKESTIILGLKEBELREKCLANSYVDKVELKMTSKQAIID-KIIASTIT 416
QY 461 TYNFE-----NGRVAVDTGSN-----MSEV-LPOIQETTA-----RIIFNGKOLMLV 501
DB 417 SKDPQOSQANRITIGPAAANDIAKELAKYLTEISVENINVTGLYLAKVFLKDKDSPTI 476
QY 502 EERRIAAV-NPSPDLPTTKPD-----MTLKEALKIAGFNEPNGNLQYQKODITBEPDNF 554
DB 477 NRTIYLVLSGPAKVEATKPDKMDVEIYKKSINQIF-----VDDF----- 517
QY 555 DQQTSQNIKNQOLAEANAIY--TVLDK-----IKNAMKNMILIRDRK 595
DB 518 -----LNTNLNKNNTTSLKEELRNKRIROSIEDYFDIKLANFTNSKEYPWR 563
QY 596 FHYDRNNIIVGADSVVKEAREVINSSTEGLLNLIDKIRKILSGYIVIEBDETEG-LKE 654
DB 564 LDFPSNEKAI-----EYKNNNNSSSIVNFIYKVKKNYK---VHSPKSSGKIED 608
QY 655 VINDRYDMLNISIIRQDG-KTFPIDF-----KKNYDKLP-----LYISNPNYK 695
DB 609 IVFDKVFSEKVVKNLGNAAERKRFQDFPSFLKSDSQLRQKLPQOITTNELKDKQISVNDPFLD 668
```

QY 696 VNVYATKEN-----TINPSE--NGDTSTNGIKILTI-----PSK 729
DB 669 ISSNNIRNDNPMFEYKIIIEPKDKDNAD-DKNGSLKIMVLSYNKTRFSK 716

RESULT 15

T28676
rhoptry protein - Plasmodium yoelii (fragment)

CjSpecies: Plasmodium yoelii
CjDate: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

CjAccession: T28676; A45521
RjAccession: T28676; A45521

Mol. Biochem. Parasitol. 76, 329-332, 1996
Mol. Biochem. Parasitol. 76, 329-332, 1996

A>Title: Comparison of two members of a multigene family coding for high-molecular mass
A/Reference number: Z20507; MUID:97077455; PMID:8920022

A/Accession: T28676
A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
A/Residues: 1-2401 <SIN>

A/Cross-references: UNIPROT:Q26226; UNIPARC:UPI000017B647; EMBL:U36927; NID:91041784; PI
R/Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.

Mol. Biochem. Parasitol. 42, 241-246, 1990
Mol. Biochem. Parasitol. 42, 241-246, 1990

A>Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple cd
A/Reference number: A45521; MUID:91101660; PMID:2270106

A/Accession: A45521
A/Status: preliminary
A/Molecule type: DNA

A/Residues: 2260-2401 <KEE>
A/Cross-references: UNIPARC:UPI000017B648; GB:M34281

Query Match 4.8%; Score 179.5; DB 2; Length 2401;
Best Local Similarity 19.2%; Pred. No. 0.81;
Matches 148; Conservative 131; Mismatches 294; Indels 199; Gaps 37;

QY 48 SELENPSENQYQSAIMSGFIRKVKSDERTFATSADNHTMNVDDQEVINKASNS-----103
DB 352 SHLNKIEITNNKSETIL--DIK-----YIG-----EITNELNKTLEDF 389
QY 104 -NKIRLEKRLVQIKIYOQREN-----TEKGLDFKLYWTDSONKKEVISDNILQPELK 157
DB 390 KKK---EKLGNKID--EYAKENVQLVVNSNILEIKHYNQ-----INIDIKKEKAK 439
QY 158 QKSSNRKRRSTAGTPVDRNDGIPLSLVEGVYDVVKNKRTPLSPWISNIHEKKGLT 217
DB 440 QNVYDQFEHMK-----TIP-----PNEWKYQKPSIRIKIM--DEFLSKVKNYNDPD 484
QY 218 K-YKSSPEKMTASDYSPEKTYGRIDKNVSPEARHPLVAAPYIVHVMENII--LSKN 274
DB 485 KYKKEKVE-----SEHNKFTLTKIKTEVSDE--IKKYENKFNDSKSLINETKKS 534
QY 275 EDOSTONTDSQRT-----ISKNTSTSRDANTVGVIS 307
DB 535 IEBEYONINTLKKVDYIKVCLANTNELITNCHNKQTLTKDKLANQIKITIKETNSIDKIYT 594
QY 308 AGYON-----GFTGNITTSAGFSNSNSTVAIDHSLSLA-----GEET 345
DB 595 DKFENILTDKTELETFTGELSLNNHESNNKELLTFYDLKANILGKNKENMLYKOFNEKE 654
QY 346 WA-ETMGATADTARLANANIRYNTGAPRYNVLPTTSLVLSKN-----QTLATIKAKE 398
DB 655 KAVEDIKKKNVDINKIVSNIEI--TIYTSIYNINEDTENIIGKSIELANTKVLKAVANY 712
QY 399 NOLSOI-----LAPNNYPSKRLADIALNAODFSSTPTIMNYYNQFLELEKT 445
DB 713 TNLNEIKELKADYDPODFGKELIKITPDEN---KIKNDIDTANQKIDKSIETLLEIKKN 768
QY 446 KQRLDLDQYVGNIAATYFENGVRVDTGSNWSVLPQIQETIARIIFN--GKDLNVERR 504
DB 769 SEVHI--DEIKGIDKLK-----KVPNKTMFNEDEPREIKELIENIVEKIDKKKNYKEI 820
QY 505 IAAVNPSPLETTKPMTLKALKIAGFNPNGNTQYQKQDITFPDFNPDQOSTONIKN 564

DB 821 DKLINEISKIEN---DKTSLEKJK-----NINLSY--GSLGNL---FLQOIDEKKK 865
QY 565 QLAELNATNIY--TVLDKIR-----LNAKNILIRDR--FHYDRNNTAVGAD-----609
DB 866 AEHTIKAMEAYIDDLNLIKKSQOEIEKENNINNDIDMDIHKEKALNISHDDYKIIYHTTS 925
QY 610 -----SVKKAHREVINSSTEGELLNIDKDIRKILSGYIVIEPDE--GLKEVINDRYD 661
DB 926 KNHEEKISDIRKNSLKITIDFSEBSYIN--DIKKELEKNVLESQNNNDINDINOYLSKIEN 982
QY 662 MLNISLRQDGYTFIDFKKYNDKLPYISNPYKVVAVYATKENTLINPSEN 713
DB 983 IYNILKLNKIKKIDKVKKEYTDE---IKNNKKIAELSNSEKIIITOLKEN 1030

Search completed: August 11, 2006, 20:26:55
Job time : 51 secs

This Page Blank (up)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2006, 20:16:47 ; Search time 204 Seconds
(without alignment)

1647.324 Million cell updates/sec

Title: US-10-780-250-1

Sequence: 1 EVKQENRLNSESSESSQGL.....TSTNGIKILIRSKQYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext: 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq.8:*
1: geneseqp1980a:*
2: geneseqp1990a:*
3: geneseqp2000a:*
4: geneseqp2001a:*
5: geneseqp2002a:*
6: geneseqp2003a:*
7: geneseqp2003b:*
8: geneseqp2004a:*
9: geneseqp2005a:*
10: geneseqp2006a:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3770	100.0	735	8 ADU91749	ADU91749 Recombina
2	3634	96.4	735	5 AAE18289	AAE18289 Bacillus
3	3634	96.4	735	8 ADI16348	ADI16348 Bacillus
4	3634	96.4	735	8 ADI16344	ADI16344 Bacillus
5	3634	96.4	735	8 ADI16344	ADI16344 Bacillus
6	3634	96.4	735	8 ADI16344	ADI16344 Bacillus
7	3634	96.4	735	8 ADI16344	ADI16344 Bacillus
8	3634	96.4	735	10 AEF05331	AEF05331 Recombina
9	3634	96.4	735	10 AEF05331	AEF05331 Wild type
10	3634	96.4	735	6 AAE35717	AAE35717 Bacillus
11	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
12	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
13	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
14	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
15	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
16	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
17	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
18	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
19	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
20	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
21	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
22	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac
23	3634	96.4	735	8 ADI19301	ADI19301 B. anthrac

24	3631	96.3	735	9 AEA52583	AEA52583 Anthrax p
25	3631	96.3	736	3 AAY56959	AAY56959 B. anthra
26	3631	96.3	763	3 AAY56960	AAY56960 B. anthra
27	3631	96.3	764	3 AAY56958	AAY56958 B. anthra
28	3631	96.3	764	4 AAB47306	AAB47306 Wild type
29	3631	96.3	764	4 AAD172869	AD172869 Anthrax v
30	3631	96.3	764	9 AEB20798	AEB20798 Bacillus
31	3631	96.3	857	7 ADE65872	ADE65872 Bacillus
32	3631	96.3	868	9 AEB20803	AEB20803 Bacillus
33	3630	96.3	735	9 ADW05492	ADW05492 B. anthra
34	3630	96.3	735	9 AEB63718	AEB63718 Protein s
35	3630	96.3	764	8 ADM12704	ADM12704 Anthrax p
36	3630	96.3	764	10 AEF01426	AEF01426 Anthrax p
37	3630	96.3	764	10 AEF06483	AEF06483 Humanized
38	3629	96.3	764	9 AED17620	AED17620 Anthrax t
39	3628	96.2	735	9 ADZ51336	ADZ51336 Anthro ac1
40	3628	96.2	764	9 AED17616	AED17616 Anthrax t
41	3627	96.2	735	5 AAMS1490	AAMS1490 Anthrax p
42	3627	96.2	735	5 AAMS1487	AAMS1487 Anthrax p
43	3627	96.2	735	9 ADM05466	ADM05466 B. anthra
44	3627	96.2	735	9 ADM05469	ADM05469 B. anthra
45	3626	96.2	735	5 AAMS1489	AAMS1489 Anthrax p

ALIGNMENTS

RESULT 1	ADU91749	standard; protein; 735 AA.
XX	ADU91749;	
AC	ADU91749;	
XX	ADU91749;	
DT	10-FEB-2005 (first entry)	
XX	10-FEB-2005 (first entry)	
DE	Recombinant PA-I protein for inhibiting anthrax toxin.	
XX	Recombinant PA-I protein for inhibiting anthrax toxin.	
KW	antibacterial; protein engineering; anthrax toxin; pore formation;	
KW	Bacillus anthracis infection.	
XX	Bacillus anthracis.	
OS	Synthetic.	
XX	US2004235136-A1.	
PN	US2004235136-A1.	
XX	25-NOV-2004.	
PD	25-NOV-2004.	
XX	17-FEB-2004; 2004US-00780250.	
PP	17-FEB-2004; 2004US-00780250.	
XX	29-MAR-2001; 2001US-00821348.	
PR	29-MAR-2001; 2001US-00821348.	
XX	(SING/) SING Y.	
PA	(KHAN/) KHANNA H.	
XX	Singh Y, Khanna H;	
PI	Singh Y, Khanna H;	
XX	WPI; 2004-821325/81.	
DR	WPI; 2004-821325/81.	
XX	New recombinant protein designated PA-I, useful for inhibiting anthrax	
PT	toxin or for treating anthrax infection in a mammal, e.g. humans.	
XX	Claim 1; SEQ ID NO 1; 13pp; English.	
PS	Claim 1; SEQ ID NO 1; 13pp; English.	
XX	The invention relates to a recombinant protein designated PA-I comprising	
CC	735 amino acids and useful for inhibiting anthrax toxin. The recombinant	
XX	protein PA-I completely inhibits the toxicity of anthrax lethal toxin and	
CC	also inhibits pore formation by native PA in cells. The recombinant	
XX	protein designated PA-I is useful for inhibiting anthrax toxin. The	
CC	composition is useful for inhibiting anthrax toxin. The recombinant	
XX	protein is also useful for treating anthrax infection in a subject. This	
CC	sequence corresponds to the amino acid sequence of the recombinant PA-I	
XX	protein of the invention.	

SQ Sequence 735 AA;
 Query Match 100.0%; Score 3770; DB 8; Length 735;
 Best Local Similarity 100.0%; Pred. No. 1.5e-226;
 Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 EVKQENRLNLSSESSOGILGYFSDLNFOAPMVTSTTGDLSPSSELENIPEBNQYF 60
DB 1 EVKQENRLNLSSESSOGILGYFSDLNFOAPMVTSTTGDLSPSSELENIPEBNQYF 60
QY 61 OSATWSGFIKVKKSDEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 OSATWSGFIKVKKSDEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNLQLPELKQKSSNSRKRSTASAGPTVPDRDN 180
DB 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNLQLPELKQKSSNSRKRSTASAGPTVPDRDN 180
QY 181 DGIPDSLEVEGYTVVKNKRFTPLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYSPEKYT 240
DB 181 DGIPDSLEVEGYTVVKNKRFTPLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYSPEKYT 240
QY 241 GRIDKNVSPPEARHPVAAVPIVHVMENIILSKNEDOSTONTDSQTRISKNTSRDAN 300
DB 241 GRIDKNVSPPEARHPVAAVPIVHVMENIILSKNEDOSTONTDSQTRISKNTSRDAN 300
QY 301 TVGVISISAGYONGFPGNTTTSAGFSNSSTVAIDHSISLAGERTWAETMGINTADTARL 360
DB 301 TVGVISISAGYONGFPGNTTTSAGFSNSSTVAIDHSISLAGERTWAETMGINTADTARL 360
QY 361 NANIRYVNTGTAPINVLPTTSLVIGKQOTLATIRAKENQLSQIILAPNNYPSKULAPIA 420
DB 361 NANIRYVNTGTAPINVLPTTSLVIGKQOTLATIRAKENQLSQIILAPNNYPSKULAPIA 420
QY 421 LNAQDDFSTPTTMYNOFLBELKTKQLRLDQOYGNATYVNFENGVRVDTGSNWSRV 480
DB 421 LNAQDDFSTPTTMYNOFLBELKTKQLRLDQOYGNATYVNFENGVRVDTGSNWSRV 480
QY 481 LPQIOETTARITFNGKDLNVERRIAANVPSDPLETTEKDMTLKALAFGNEPENGUL 540
DB 481 LPQIOETTARITFNGKDLNVERRIAANVPSDPLETTEKDMTLKALAFGNEPENGUL 540
QY 541 OYQGGDITPEDPNPOOTSQNTKQNLAEINATNITVLDKILNKNMILINDKXPHYDR 600
DB 541 OYQGGDITPEDPNPOOTSQNTKQNLAEINATNITVLDKILNKNMILINDKXPHYDR 600
QY 601 NNIAVGADESYYKAAHREVINSSTEGLLNIDKDIRKILSGYIVELBDETEGLKAVIINDRY 660
DB 601 NNIAVGADESYYKAAHREVINSSTEGLLNIDKDIRKILSGYIVELBDETEGLKAVIINDRY 660
QY 661 DMLNITSSLRQDQKTIFDFKKNYDKLPLYSNPYKVNVAATKENTIIINPSKNGTSTNG 720
DB 661 DMLNITSSLRQDQKTIFDFKKNYDKLPLYSNPYKVNVAATKENTIIINPSKNGTSTNG 720
QY 721 IKKILIPSKKGYEIG 735
DB 721 IKKILIPSKKGYEIG 735
  
```

RESULT 2
 AAE18289 standard; protein; 735 AA.
 AAE18289;
 07-MAY-2002 (first entry)
 Bacillus subtilis protective antigen (PA) domain 1+2+3+4.
 Immunogenic reagent; immune response; protective antigen; PA; vaccine;
 Bacillus anthracis infection; antibacterial.
 Bacillus subtilis.

XX Key Location/Qualifiers
 FH Misc-difference 285
 FT /note= "Encoded by GAA"
 XX
 XX WO200204646-A1.
 XX
 XX 17-JAN-2002.
 XX
 XX 06-JUL-2001; 2001WO-GB003065.
 XX
 XX 08-JUL-2000; 2000GB-00016702.
 XX
 XX (MINA) UK SEC FOR DEFENCE.
 XX
 XX Williamson ED, Miller J, Walker NJ, Baillie LMJ, Holden PT;
 XX Flick-Smith HC, Bullifent HU, Tibball RW, Topping AW;
 XX
 XX WPI: 2002-171720/22.
 XX
 XX N-PSDB; AAD29121.
 XX
 XX New immunogenic reagent having a polypeptide of the full length
 XX PT Protective Antigen of Bacillus anthracis, useful for treating B.
 XX PT anthracis infection or in preparing a medicament for the prophylaxis or
 XX PT treatment of the infection.
 XX
 XX Example 5; Fig 3; 40pp; English.
 XX
 XX The present invention relates to an immunogenic reagent, which produces
 XX CC an immune response that is protective against Bacillus anthracis. The
 XX CC reagent comprises one or more polypeptides which together represent up to
 XX CC three domains of the full length Protective Antigen (PA) of Bacillus
 XX CC anthracis or variants of these, and at least one of the domains comprises
 XX CC domain 1 or domain 4 of PA or its variant. The invention is used as a
 XX CC vaccine. The immunogenic reagent is useful in the preparation of a
 XX CC medicament for the prophylaxis or treatment of B. anthracis infection.
 XX CC The present sequence is Bacillus subtilis protective antigen domain
 XX CC 1+2+3+4
 XX
 SQ Sequence 735 AA;
 Query Match 96.4%; Score 3634; DB 5; Length 735;
 Best Local Similarity 97.2%; Pred. No. 4.6e-218;
 Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;

```

QY 1 EVKQENRLNLSSESSOGILGYFSDLNFOAPMVTSTTGDLSPSSELENIPEBNQYF 60
DB 1 EVKQENRLNLSSESSOGILGYFSDLNFOAPMVTSTTGDLSPSSELENIPEBNQYF 60
QY 61 OSATWSGFIKVKKSDEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 OSATWSGFIKVKKSDEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNLQLPELKQKSSNSRKRSTASAGPTVPDRDN 180
DB 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNLQLPELKQKSSNSRKRSTASAGPTVPDRDN 180
QY 181 DGIPDSLEVEGYTVVKNKRFTPLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYSPEKYT 240
DB 181 DGIPDSLEVEGYTVVKNKRFTPLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYSPEKYT 240
QY 241 GRIDKNVSPPEARHPVAAVPIVHVMENIILSKNEDOSTONTDSQTRISKNTSRDAN 300
DB 241 GRIDKNVSPPEARHPVAAVPIVHVMENIILSKNEDOSTONTDSQTRISKNTSRDAN 300
QY 301 TVGVISISAGYONGFPGNTTTSAGFSNSSTVAIDHSISLAGERTWAETMGINTADTARL 360
DB 301 TVGVISISAGYONGFPGNTTTSAGFSNSSTVAIDHSISLAGERTWAETMGINTADTARL 360
QY 361 NANIRYVNTGTAPINVLPTTSLVIGKQOTLATIRAKENQLSQIILAPNNYPSKULAPIA 417
DB 361 NANIRYVNTGTAPINVLPTTSLVIGKQOTLATIRAKENQLSQIILAPNNYPSKULAPIA 417
  
```



```

QY 418 PIALNAQDDSSPTITMANNQFLEKTKQRLDQVYGNIAATYFENGRAVVDGSMW 477
DB 418 PIALNAQDDSSPTITMANNQFLEKTKQRLDQVYGNIAATYFENGRAVVDGSMW 477
QY 478 SEVLPQIOETTAIIIFNGKDLNVERRIAANVPSDLETTKPDMTLKEALKAIFGNEPN 537
DB 478 SEVLPQIOETTAIIIFNGKDLNVERRIAANVPSDLETTKPDMTLKEALKAIFGNEPN 537
QY 538 GNLQYQKDIITEFDNFDOQTSQINQOLAELNATNIYTVLDKIKLNAKNNILIRKRFH 597
DB 538 GNLQYQKDIITEFDNFDOQTSQINQOLAELNATNIYTVLDKIKLNAKNNILIRKRFH 597
QY 598 YDRNNIAGADESVYKEAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVIN 657
DB 598 YDRNNIAGADESVYKEAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVIN 657
QY 658 DRYDMLNISLRODGTFTIDFKKNDKLPYISNPYKVVAVTENTIIINSENGDTS 717
DB 658 DRYDMLNISLRODGTFTIDFKKNDKLPYISNPYKVVAVTENTIIINSENGDTS 717
QY 718 TNGIKKILIFSKKGYEIG 735
DB 718 TNGIKKILIFSKKGYEIG 735

RESULT 3
ADL16348 standard; protein; 735 AA.
XX ADL16348;
AC
XX
DT 06-MAY-2004 (first entry)
XX
DE Bacillus anthracis PA protein from pBP105 vector.
XX
KW Immunogenic; vaccine; lethal infection; protective antigen; PA;
XX lethal factor; LF; anthrax; antibacterial.
XX
OS Bacillus anthracis.
XX
PN US2004028695-A1.
XX
PD 12-FEB-2004.
XX
PF 28-MAR-2003; 2003US-00402466.
XX
PR 12-APR-2002; 2002US-0372152P.
XX
PA (PARK/) PARK S.
XX (GIRI/) GIRI L.
XX
PI Park S, Giri L;
XX
DR WPI; 2004-168865/16.
XX N-PSDB; ADL16346.
XX
PT Immunogenic composition for preparing a vaccine against a lethal
PT infection of Bacillus anthracis in an animal, comprises an immunizing
PT amount of a recombinant B. anthracis protective antigen protein and/or
PT lethal factor protein.
XX
PS Disclosure; SEQ ID NO 13; 143pp; English.
XX
XX The present invention relates to immunogenic compositions useful for
XX preparing a vaccine against a lethal infection of Bacillus anthracis in
XX an animal. The compositions comprise an immunizing amount of a
XX recombinant B. anthracis protective antigen (rPA) protein and a
XX recombinant B. anthracis lethal factor (rLF) protein. The composition and
XX methods are useful in protecting against anthrax or lethal infections
XX caused by B. anthracis. The present sequence is Bacillus anthracis PA
XX protein from pBP105 vector used in the invention.
XX
SQ Sequence 735 AA;

```

```

Query Match 96.4%; Score 3634; DB 8; Length 735;
Best Local Similarity 97.2%; Pred. No. 4,6e-218;
Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;

QY 1 EVKQENRLINSESSSSQGLIGYFSDNFOAPMVVTSSTTGDLISIPSELENIIPSENQYF 60
DB 1 EVKQENRLINSESSSSQGLIGYFSDNFOAPMVVTSSTTGDLISIPSELENIIPSENQYF 60
QY 61 QSAIWSGPIKVKKSDRYTATSADNHVTWVDDQEVYINAKSNKRLKRGRIYQIKQY 120
DB 61 QSAIWSGPIKVKKSDRYTATSADNHVTWVDDQEVYINAKSNKRLKRGRIYQIKQY 120
QY 121 QRENPTKGLDPLVYTDSONKKEVSSDNLOPELKOKSNSRKRSTAGPTVDRDN 180
DB 121 QRENPTKGLDPLVYTDSONKKEVSSDNLOPELKOKSNSRKRSTAGPTVDRDN 180
QY 181 DGIPDGLVEGYTVDVKNRTPFLSPWISNIEKKGITTKYKSSPEKMTASDPYSDEKYV 240
DB 181 DGIPDGLVEGYTVDVKNRTPFLSPWISNIEKKGITTKYKSSPEKMTASDPYSDEKYV 240
QY 241 GRIDKNVSPKARRPLVAAVPIVHVDMENTIISKNEQSTQNTDSQTRTISKNTSRDAN 300
DB 241 GRIDKNVSPKARRPLVAAVPIVHVDMENTIISKNEQSTQNTDSQTRTISKNTSRDAN 300
QY 301 TGVGVSISAGYQNGF--TGNIITTSAGFSNSNSTVAIDHSLSIAGERTVAETMTADT 357
DB 301 TSEVHGNAEVAHSFDFDIGSV--SAGFSNSNSTVAIDHSLSIAGERTVAETMTADT 357
QY 358 ARLNANIRVYNTGTAPITYVLPPTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLA 417
DB 358 ARLNANIRVYNTGTAPITYVLPPTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLA 417
QY 418 PIALNAQDDSSPTITMANNQFLEKTKQRLDQVYGNIAATYFENGRAVVDGSMW 477
DB 418 PIALNAQDDSSPTITMANNQFLEKTKQRLDQVYGNIAATYFENGRAVVDGSMW 477
QY 478 SEVLPQIOETTAIIIFNGKDLNVERRIAANVPSDLETTKPDMTLKEALKAIFGNEPN 537
DB 478 SEVLPQIOETTAIIIFNGKDLNVERRIAANVPSDLETTKPDMTLKEALKAIFGNEPN 537
QY 538 GNLQYQKDIITEFDNFDOQTSQINQOLAELNATNIYTVLDKIKLNAKNNILIRKRFH 597
DB 538 GNLQYQKDIITEFDNFDOQTSQINQOLAELNATNIYTVLDKIKLNAKNNILIRKRFH 597
QY 598 YDRNNIAGADESVYKEAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVIN 657
DB 598 YDRNNIAGADESVYKEAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVIN 657
QY 658 DRYDMLNISLRODGTFTIDFKKNDKLPYISNPYKVVAVTENTIIINSENGDTS 717
DB 658 DRYDMLNISLRODGTFTIDFKKNDKLPYISNPYKVVAVTENTIIINSENGDTS 717
QY 718 TNGIKKILIFSKKGYEIG 735
DB 718 TNGIKKILIFSKKGYEIG 735

RESULT 4
ADL16344
ID ADL16344 standard; protein; 735 AA.
XX
XX ADL16344;
AC
XX
DT 06-MAY-2004 (first entry)
XX
DE Bacillus anthracis wild-type rPA protein #1.
XX
KW Immunogenic; vaccine; lethal infection; protective antigen; PA;
XX lethal factor; LF; anthrax; antibacterial.
XX
OS Bacillus anthracis.
XX

```

PN US2004028695-A1.
 XX 12-FEB-2004.
 PD 28-MAR-2003; 2003US-00402466.
 PF 12-APR-2002; 2002US-0372152P.
 PR (PARK/) PARK S.
 XX (GIRI/) GIRI L.
 XX Park S, Giri L;
 XX WPI: 2004-168865/16.
 DR N-PEDB; ADL16343, ADL16345.
 CC Immunogenic composition for preparing a vaccine against a lethal
 PT infection of *Bacillus anthracis* in an animal, comprises an immunizing
 PT amount of a recombinant *B. anthracis* protective antigen protein and/or
 PT lethal factor protein.
 XX Claim 4; SEQ ID NO 9; 143pp; English.
 PS The present invention relates to immunogenic compositions useful for
 CC preparing a vaccine against a lethal infection of *Bacillus anthracis* in
 CC an animal. The compositions comprise an immunising amount of a
 CC recombinant *B. anthracis* protective antigen (rPA) protein and a
 CC recombinant *B. anthracis* lethal factor (rLF) protein. The composition and
 CC methods are useful in protecting against anthrax or lethal infections
 CC caused by *B. anthracis*. The present sequence is *Bacillus anthracis* wild-
 CC type rPA protein used in the invention.
 XX Sequence 735 AA;

Query Match 96.4%; Score 3634; DB 8; Length 735;
 Best Local Similarity 97.2%; Pred. No. 4,6e-218;
 Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;

QY 1 EVKQNRRLNBSSESSOGLGTYPSDLNFOAPMVYTSSTGSLSPSSLENIPEBNQYF 60
 DB 1 EVKQNRRLNBSSESSOGLGTYPSDLNFOAPMVYTSSTGSLSPSSLENIPEBNQYF 60
 QY 61 OSATWSGFIVKVKSEBYFATSDAHNVMTWVDQEVINKASNSNKLRLKEKGLYOIKIY 120
 DB 61 OSATWSGFIVKVKSEBYFATSDAHNVMTWVDQEVINKASNSNKLRLKEKGLYOIKIY 120
 QY 121 QRENPTKGLDPKLYWTDSONKKEVISSDNIQLPELKQKSSNRKRSTSGAPVPPDRN 180
 DB 121 QRENPTKGLDPKLYWTDSONKKEVISSDNIQLPELKQKSSNRKRSTSGAPVPPDRN 180
 QY 181 DGIPSLVEGGTYVVKKRTFLSPWISNIHEKGLTKYKSPKMWSTASDYSPEKKT 240
 DB 181 DGIPSLVEGGTYVVKKRTFLSPWISNIHEKGLTKYKSPKMWSTASDYSPEKKT 240
 QY 241 GRIDGNVSPPEARHPVAAPIYVHMENIISKNEDOSTONTDSORTISKYSTSRDAN 300
 DB 241 GRIDGNVSPPEARHPVAAPIYVHMENIISKNEDOSTONTDSORTISKYSTSRDAN 300
 QY 301 TVGVISAGYONGF--TGNITTSAGFSNSNSTVAIDHSLSLAGEERTVAETGINTADT 357
 DB 301 TSEVHGNAEVAHSPFDIGSV--SAGFSNSNSTVAIDHSLSLAGEERTVAETGINTADT 357
 QY 358 ARLANINIRYVNTGTAPIYVLTTSIVLAKNOTLATIRAKENQSLAIPNNYYSKTLA 417
 DB 358 ARLANINIRYVNTGTAPIYVLTTSIVLAKNOTLATIRAKENQSLAIPNNYYSKTLA 417
 QY 418 PLTANAODFSSSTPTMNNYNOFLELEKTKQLRLDTPDOVGNATATNFENGRVAVTGSNW 477
 DB 418 PLTANAODFSSSTPTMNNYNOFLELEKTKQLRLDTPDOVGNATATNFENGRVAVTGSNW 477
 QY 478 SEVLFOIETTARIIFNGKDLNLVERRIAANVPSDELTTKPDMLTKEALKIAFGNEEN 537
 DB 478 SEVLFOIETTARIIFNGKDLNLVERRIAANVPSDELTTKPDMLTKEALKIAFGNEEN 537

QY 538 GNLOYOGKDITTEPDPNPQOOTSQNIKNQOLAEANATNITYVLDPKIKANAKNILLRDKRFH 597
 DB 538 GNLOYOGKDITTEPDPNPQOOTSQNIKNQOLAEANATNITYVLDPKIKANAKNILLRDKRFH 597
 QY 598 YDRNNIAVGADESIVKKAHREVINSSTEGLLINIDKDIRKLISGYIVIEDTEGLKEVIN 657
 DB 598 YDRNNIAVGADESIVKKAHREVINSSTEGLLINIDKDIRKLISGYIVIEDTEGLKEVIN 657
 QY 658 DRYDMLNITSSLRQDGKTPIIDPKTNDKLPYISNPNYVNNYATVKENTIIINPSNGQTS 717
 DB 658 DRYDMLNITSSLRQDGKTPIIDPKTNDKLPYISNPNYVNNYATVKENTIIINPSNGQTS 717
 QY 718 TNGIKKILIFSKKGYEIG 735
 DB 718 TNGIKKILIFSKKGYEIG 735
 RESULT 5
 ADR40460
 ID ADR40460 standard; protein; 735 AA.
 AC ADR40460;
 DT 18-NOV-2004 (first entry)
 XX *Bacillus anthracis* protective antigen F313A F314A mutant.
 XX antibacterial; vaccine; protective antigen; PA; proteolytic degradation;
 XX *B. anthracis* toxin; *B. anthracis* infection; passive immunisation;
 XX inhalation anthrax; cutaneous anthrax; gastrointestinal anthrax; mutant;
 XX murein.
 XX *Bacillus anthracis*.
 OS Synthetic.
 OS Key Location/Qualifiers
 FH Misc-difference 313..314
 FT /note= "Wild type Phe substituted by Ala"
 PN US2004171121-A1.
 PD 02-SEP-2004.
 XX 08-AUG-2003; 2003US-00638006.
 PF 09-AUG-2002; 2002US-0402285P.
 PR (LEPP/) LEPPILA S H.
 XX (ROSO/) ROBOVITZ M J.
 PA (HSUS/) HSU S D.
 XX Leppila SH, Rosovitz MJ, Hsu SD;
 PI WPI: 2004-625107/60.
 DR Novel *Bacillus anthracis* protective antigen having mutations conferring
 PT enhanced resistance to proteolytic degradation compared to wild type
 PT antigen, useful for inducing antibodies having neutralizing activity
 PT anthrax toxin.
 PS Claim 11; Page: 30pp; English.
 XX The invention describes a recombinant *Bacillus anthracis* protective
 CC antigen (PA), modified to incorporate one or more mutations comprising an
 CC amino acid deletion or substitution in a flexible, exposed, or loop
 CC segment of the PA protein, where the one or more mutations confers
 CC enhanced resistance to proteolytic degradation compared to wild type PA.
 CC A pharmaceutical composition (I) comprising PA and physiologically
 CC acceptable carrier or a composition (II) comprising a synthetic construct
 CC encoding PA and a carrier is useful for inducing serum antibodies that
 CC have neutralizing activity for a *B. anthracis* toxin which involves
 CC administering (I) to a mammal (human), sufficient to elicit production of

CC the antibodies. The antibodies protect the mammal against the infection.
 CC (I) is useful for vaccinating a human against B.anthraxis infection. An
 CC anti-PA-antibody containing composition is useful for passively
 CC immunising a mammal against the toxic effect of B.anthraxis. PA is useful
 CC as vaccines to induce serum antibodies which are useful to prevent, treat
 CC or reduced the severity of infections caused by B.anthraxis, such as
 CC inhalation anthrax, cutaneous anthrax and/or gastrointestinal anthrax. PA
 CC exhibits enhanced stability to proteolytic degradation that is increased
 CC by at least 25% compared to proteolytic degradation of wild-type PA under
 CC comparable conditions. This is the amino acid sequence of protective
 CC antigen (PA) F313A F314A mutant. Note: This sequence does not appear in
 CC the specification but has been created by the indexer using information
 CC given in the invention.

XX Sequence 735 AA;

Query Match 96.4%; Score 3634; DB 8; Length 735;

Best Local Similarity 96.8%; Pred. No. 4.6e-218;

Matches 716; Conservative 2; Mismatches 12; Indels 10; Gaps 2;

QY 1 EYKQENRLINSESSSQGLGYFSDLNFOAPMVVTSSTTGDLSTPSSELENIPESENQYF 60
 DB 1 EYKQENRLINSESSSQGLGYFSDLNFOAPMVVTSSTTGDLSTPSSELENIPESENQYF 60
 QY 61 QSAIWSGFIKVKKSDYTFATSDAHNVMTWVDDQEVINKASNSNKLRLKGRLYQIKIQY 120
 DB 61 QSAIWSGFIKVKKSDYTFATSDAHNVMTWVDDQEVINKASNSNKLRLKGRLYQIKIQY 120
 QY 121 QRENPEFKGIDFLYNTDSQNKKEVISDNLOPELKQKSSNRKRGSTAGTVPDRDN 180
 DB 121 QRENPEFKGIDFLYNTDSQNKKEVISDNLOPELKQKSSNRKRGSTAGTVPDRDN 180
 QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSPKKSTASDPYDFEKT 240
 DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSPKKSTASDPYDFEKT 240
 QY 241 GRIDKVSPEARHPVLAAPYIVHVDMENTILSKNEQOSTONTOSQRTTSKNTSTSRDN 300
 DB 241 GRIDKVSPEARHPVLAAPYIVHVDMENTILSKNEQOSTONTOSQRTTSKNTSTSRDN 300
 QY 301 T-----VGVISISAGYONGFTGNITTSAGFSNSNSTVAIDHSLSLAGEBTWMTGLNTA 355
 DB 301 SEVHGNAEYHAASADTGG-----SVSAGFSNSNSTVAIDHSLSLAGEBTWMTGLNTA 355
 QY 356 DPARLANIRYVNTGTAPYVNLPTTSLVYGKQQTATIKAKENQULSLAPNNYPSKX 415
 DB 356 DPARLANIRYVNTGTAPYVNLPTTSLVYGKQQTATIKAKENQULSLAPNNYPSKX 415
 QY 416 LAPIALNADDDPSSTITMNTNOFLELEKTKQRLDTDOYGGIATYFNENGVRVDG 475
 DB 416 LAPIALNADDDPSSTITMNTNOFLELEKTKQRLDTDOYGGIATYFNENGVRVDG 475
 QY 476 NMSEVLPJOIETTARIIFNGKDLNVERRIAAVNPSPDETTEPDMTLKEALTIARFNE 535
 DB 476 NMSEVLPJOIETTARIIFNGKDLNVERRIAAVNPSPDETTEPDMTLKEALTIARFNE 535
 QY 536 PNGNTQYQSKDITTEPDPNDOQTSQNIKQQLAELNANTYTVLDTIKLANKNMILLRDK 595
 DB 536 PNGNTQYQSKDITTEPDPNDOQTSQNIKQQLAELNANTYTVLDTIKLANKNMILLRDK 595
 QY 596 FHYDRNNIAGADESVYKGAHREVINSSTREGLLANTDKDIRKLSGYIEIETBGLAKY 655
 DB 596 FHYDRNNIAGADESVYKGAHREVINSSTREGLLANTDKDIRKLSGYIEIETBGLAKY 655
 QY 656 INDRYDMLNISSLRODGKTFIDPKKYNNDKPLPTISNPYKVNYYATKENTIIINPSENG 715
 DB 656 INDRYDMLNISSLRODGKTFIDPKKYNNDKPLPTISNPYKVNYYATKENTIIINPSENG 715
 QY 716 TSTNGIKILLIPSKGYEIG 735
 DB 716 TSTNGIKILLIPSKGYEIG 735

RESULT 6

ID ADR40458
 ADR40458 standard; protein; 735 AA.

AC ADR40458;

DT 18-NOV-2004 (first entry)

DE Bacillus anthracis wild type protective antigen seqid 2.

XX antibacterial; vaccine; protective antigen; PA; proteolytic degradation;

KW B.anthraxis toxin; B.anthraxis infection; passive immunisation;

XX inhalation anthrax; cutaneous anthrax; gastrointestinal anthrax.

OS Bacillus anthracis.

PN US2004171121-A1.

PD 02-SEP-2004.

XX 08-AUG-2003; 2003US-00638006.

PR 09-AUG-2002; 2002US-0402285P.

XX (LEPP/) LEPLA S H.

PA (ROSO/) ROSOVITZ M J.

XX (HSUS/) HSU S D.

PI Leppla SH, Roseovitz MJ, Hsu SD;

DR WPI; 2004-625107/60.

PT Novel Bacillus anthracis protective antigen having mutations conferring

PT enhanced resistance to proteolytic degradation compared to wild type

PT anthrax, useful for inducing antibodies having neutralizing activity

PT anthrax toxin.

PS Example 5; SEQ ID NO 2; 30pp; English.

XX The invention describes a recombinant Bacillus anthracis protective

CC antigen (PA), modified to incorporate one or more mutations comprising an

CC amino acid deletion or substitution in a flexible, exposed, or loop

CC segment of the PA protein, where the one or more mutations confers

CC enhanced resistance to proteolytic degradation compared to wild type PA.

CC A pharmaceutical composition (i) comprising PA and physiologically

CC acceptable carrier or a composition (ii) comprising a synthetic construct

CC encoding PA and a carrier is useful for inducing serum antibodies that

CC have neutralising activity for a B. anthracis toxin which involves

CC administering (i) to a mammal (human), sufficient to elicit production of

CC the antibodies. The antibodies protect the mammal against the infection.

CC (I) is useful for vaccinating a human against B.anthraxis infection. An

CC anti-PA-antibody containing composition is useful for passively

CC immunising a mammal against the toxic effect of B.anthraxis. PA is useful

CC as vaccines to induce serum antibodies which are useful to prevent, treat

CC or reduced the severity of infections caused by B.anthraxis, such as

CC inhalation anthrax, cutaneous anthrax and/or gastrointestinal anthrax. PA

CC exhibits enhanced stability to proteolytic degradation that is increased

CC by at least 25% compared to proteolytic degradation of wild-type PA under

CC comparable conditions. This is the amino acid sequence of wild type

XX protective antigen (PA).

XX Sequence 735 AA;

Query Match 96.4%; Score 3634; DB 8; Length 735;

Best Local Similarity 97.2%; Pred. No. 4.6e-218;

Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;

QY 1 EYKQENRLINSESSSQGLGYFSDLNFOAPMVVTSSTTGDLSTPSSELENIPESENQYF 60
 DB 1 EYKQENRLINSESSSQGLGYFSDLNFOAPMVVTSSTTGDLSTPSSELENIPESENQYF 60
 QY 61 QSAIWSGFIKVKKSDYTFATSDAHNVMTWVDDQEVINKASNSNKLRLKGRLYQIKIQY 120
 DB 61 QSAIWSGFIKVKKSDYTFATSDAHNVMTWVDDQEVINKASNSNKLRLKGRLYQIKIQY 120

```

Db 61 OSAINSGEPIKVKKSDDEYTFATSDAHNVHTMWDDQEVINKANSNKKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDPLKLYWTDSONKKEVYSSDNILQPELKOKSSNRKKRSTAGPTVPPDRN 180
Db 121 QRENPTKGLDPLKLYWTDSONKKEVYSSDNILQPELKOKSSNRKKRSTAGPTVPPDRN 180
QY 181 DGIPLSLVEGYTVVKNKRFTLSPWISNIHEKGLTKYKSSPEKMTASDYSDFEKT 240
Db 181 DGIPLSLVEGYTVVKNKRFTLSPWISNIHEKGLTKYKSSPEKMTASDYSDFEKT 240
QY 241 GRIDKNVSPKARHPLVAAYPIVHVMENIILSKNEDOSTONTDSQTRTISKYTSRDN 300
Db 241 GRIDKNVSPKARHPLVAAYPIVHVMENIILSKNEDOSTONTDSQTRTISKYTSRDN 300
QY 301 TVGVISAGYGNF--TGNITTSAGFSNSNSTVAIDHSLSLAGERTWAEFTMGINTADT 357
Db 301 TVGVISAGYGNF--TGNITTSAGFSNSNSTVAIDHSLSLAGERTWAEFTMGINTADT 357
QY 358 ARINANIRVNTGTPIYVNLPTTSLVYGKNOTLATITKAKENQSLQILAPNNYPSKNLA 417
Db 358 ARINANIRVNTGTPIYVNLPTTSLVYGKNOTLATITKAKENQSLQILAPNNYPSKNLA 417
QY 418 PIALNAODFSSPTITMYNNOGLELEKTKQLRLDTPQVYGNATYTFENGRRVDTGSNW 477
Db 418 PIALNAODFSSPTITMYNNOGLELEKTKQLRLDTPQVYGNATYTFENGRRVDTGSNW 477
QY 478 SEVLPQIETTARIIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKTAFGENEPN 537
Db 478 SEVLPQIETTARIIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKTAFGENEPN 537
QY 538 GNLQVOGKDITFEDNPDQOTSQNKQLAELNATNYVLKIKLANAMNLLIDKRP 597
Db 538 GNLQVOGKDITFEDNPDQOTSQNKQLAELNATNYVLKIKLANAMNLLIDKRP 597
QY 598 YDRNNIANGADESVYKARREVINSTEGLINIDKIRKILSGYIVLEIDTEGLKEVIN 657
Db 598 YDRNNIANGADESVYKARREVINSTEGLINIDKIRKILSGYIVLEIDTEGLKEVIN 657
QY 658 DRYDMNLNSSLRQDGKTFIDPKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTS 717
Db 658 DRYDMNLNSSLRQDGKTFIDPKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTS 717
QY 718 TNGIKKILIFSKGYEIG 735
Db 718 TNGIKKILIFSKGYEIG 735

RESULT 7
ADV90293
1D ADV90293 standard, protein, 735 AA.
XX
XX ADV90293;
XX
XX
XX 10-MAR-2005 (first entry)
DE Protease-hydrolysed polypeptide #70.
XX
XX Protease; immune disorder; inflammation; musculoskeletal disease;
XX dermatological disease; gastrointestinal disease; endocrine disease;
XX metabolic disorder; cancer; hematological disease;
XX cardiovascular disease; neurological disease; neurodegenerative disease;
XX growth disorder; respiratory disease; genitourinary disease;
XX gynecological disorder; nutritional disorder; infection; cytostatic;
XX gastroenterical-gen./antiinflammatory; antiaesthetic; analgesic;
XX antiarthritic; osteopathic; antidiabetic; nephrotoxic;
XX cardiovascular-gen./immunopreservative; respiratory-gen.; antipneumatic;
XX antiallergic; dermatological; enzyme; hydrolysis.
XX
XX Homo sapiens.
XX
XX WO2004113522-A1.
XX
XX 29-DEC-2004.

```

```

XX 18-JUN-2004; 2004WO-BP051173.
PF 18-JUN-2003; 2003EP-00013819.
XX 18-JUN-2003; 2003EP-00013819.
PR 10-NOV-2003; 2003EP-00025851.
XX 10-NOV-2003; 2003EP-00025851.
PR 11-FEB-2004; 2004EP-00003058.
XX 11-FEB-2004; 2004EP-00003058.
PA (DIRE) DIREVO BIOTECH AG.
XX
PI Haupts U, Koltermann A, Scheidig A, Voetemeier C, Ketting U;
DR WPI; 2005-057985/06.
XX
PT Proteases with defined specificity for a target substrate useful for
PT treating a specific disease related to the target substrate, such as
PT cancer, asthma, diabetes, inflammatory disorders and psoriasis.
XX
XX
XX Claim 38; SEQ ID NO 123; 250pp; English.
XX
PS The invention relates to the use of a protease with defined specificity
CC for a target substrate for preparing a medicament for the treatment of a
CC specific disease related to the target substrate. The invention also
CC relates to a pharmaceutical or diagnostic composition comprising one or
CC more enzymes in the use cited, optionally comprising pharmaceutically or
CC diagnostically acceptable carriers, excipients and/or auxiliary agents, a
CC method for cleaving a target substrate in vivo or in vitro comprising
CC contacting the target substrate with a protease as cited in the use
CC mentioned, and a method for treatment of a disease in a patient connected
CC with a specific target substrate comprising administering to the patient
CC a protease with defined specificity for the specific target substrate.
CC The protease hydrolyzes the target substrate and eliminates or reduces
CC one or more biological activities, physico-chemical properties or
CC pharmacological properties of the target protein and/or activates or
CC increases one or more biological activities, physico-chemical properties
CC or pharmacological properties of the target protein, and/or adds one or
CC more biological activities, physico-chemical properties or
CC pharmacological properties to the target protein. The protease may be
CC administered to treat immune disorders, inflammatory disorders,
CC musculoskeletal diseases, dermatological diseases, gastrointestinal
CC diseases, endocrine diseases, metabolic disorder, cancers, hematological
CC diseases, cardiovascular diseases, neurological diseases,
CC neurodegenerative diseases, growth disorders, respiratory diseases,
CC genitourinary diseases, gynecological disorders, nutritional disorders
CC and infections. This sequence represents a polypeptide hydrolysed by a
CC protease used in the scope of the invention.
XX
XX
XX Sequence 735 AA;
SQ
Query Match 96.4%; Score 3634; DB 9; Length 735;
Best Local Similarity 97.2%; Pred. No. 4, 6e-218; Indels 6; Gaps 3;
Matches 717; Conservative 4; Mismatches 11;
QY 1 EVKQENRLINSESSSSQGLGYFSDLNFOAPMVVYTSSTTGSLSPSSLENIPENOYF 60
Db 1 EVKQENRLINSESSSSQGLGYFSDLNFOAPMVVYTSSTTGSLSPSSLENIPENOYF 60
QY 61 OSAINSGEPIKVKKSDDEYTFATSDAHNVHTMWDDQEVINKANSNKKIRLEKGLYQIKIY 120
Db 61 OSAINSGEPIKVKKSDDEYTFATSDAHNVHTMWDDQEVINKANSNKKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDPLKLYWTDSONKKEVYSSDNILQPELKOKSSNRKKRSTAGPTVPPDRN 180
Db 121 QRENPTKGLDPLKLYWTDSONKKEVYSSDNILQPELKOKSSNRKKRSTAGPTVPPDRN 180
QY 181 DGIPLSLVEGYTVVKNKRFTLSPWISNIHEKGLTKYKSSPEKMTASDYSDFEKT 240
Db 181 DGIPLSLVEGYTVVKNKRFTLSPWISNIHEKGLTKYKSSPEKMTASDYSDFEKT 240
QY 241 GRIDKNVSPKARHPLVAAYPIVHVMENIILSKNEDOSTONTDSQTRTISKYTSRDN 300
Db 241 GRIDKNVSPKARHPLVAAYPIVHVMENIILSKNEDOSTONTDSQTRTISKYTSRDN 300

```

QY	301	TVGSVISAIGYQNGF---TGNIITTSAGFSNSNSTVAIDHSLSLAGERTWAETMGTLTAOT	357
Db	300	TSEVHGNAEVAHASFPDIGSV--SAGFSNSNSTVAIDHSLSLAGERTWAETMGTLTAOT	357
QY	358	ALANINIRVYNGTAPIYVNVLPPTSVLGKNOQTATIKKAKENOLSOILAPNNYPSKNI	417
Db	358	ARLNINIRVYNGTAPIYVNVLPPTSVLGKNOQTATIKKAKENOLSOILAPNNYPSKNI	417
QY	418	PLALNAGDDFSSTPIITMANNOPLELEKTYQLRLDTQVGNLATYFENGRRVVDGSMW	477
Db	418	PLALNAGDDFSSTPIITMANNOPLELEKTYQLRLDTQVGNLATYFENGRRVVDGSMW	477
QY	478	SEVLFOIOTTARIIFNGKDLNVERRIAANVPDLETTKPMTLKEALKIAPFGENEN	537
Db	478	SEVLFOIOTTARIIFNGKDLNVERRIAANVPDLETTKPMTLKEALKIAPFGENEN	537
QY	538	GNLQYQSGKDIETFDENPDQOTSQNIKNQLAELNATNIYTVLDKIKLANKNILIRDKRFH	597
Db	538	GNLQYQSGKDIETFDENPDQOTSQNIKNQLAELNATNIYTVLDKIKLANKNILIRDKRFH	597
QY	598	YDRNNIANGAGADSVYKAEARVYINSTBGLINDIDIKILISGYVEIJEDEGLKEVIN	657
Db	598	YDRNNIANGAGADSVYKAEARVYINSTBGLINDIDIKILISGYVEIJEDEGLKEVIN	657
QY	658	DRYDMLNTSSLKQDGKTFIDFKKYNKDLPLYISNPYKVVYAVTKENTINPSENGDTS	717
Db	658	DRYDMLNTSSLKQDGKTFIDFKKYNKDLPLYISNPYKVVYAVTKENTINPSENGDTS	717
QY	718	TNGIKKILILFSKKGYEIG	735
Db	718	TNGIKKILILFSKKGYEIG	735

RESULT 8	
AEF05331	
ID	AEF05331 standard; protein; 735 AA.
XX	
XX	AEF05331;
XX	
DT	23-FEB-2006 (first entry)
XX	
DE	Recombinant Protective Antigen, rPA, SEQ ID NO:5.
XX	
KW	protective antigen; PA; antigen; Bacillus anthracis infection;
XX	antibacterial; infection.
XX	
OS	Bacillus anthracis.
XX	
PN	MO2005123764-A1.
XX	
PD	
XX	29-DEC-2005.
XX	
PF	16-JUN-2005; 2005WO-GB002360.
XX	
PR	16-JUN-2004; 2004GB-00013475.
XX	
PR	16-JUN-2004; 2004US-0579687P.
XX	
PR	15-JUN-2005; 2005US-00153865.
XX	
PA	(HEAL-) HEALTH PROTECTION AGENCY.
XX	
PA	(DYNP-) DYNPOT VACCINE CO LLC.
XX	
PI	Brehm J, Mcenteer I, Vincent P, Allison N, Brehm R, Jack G;
XX	Herbert M, Solow BT, Arroyo J, Lapcevlch RK;
XX	
DR	WPI; 2006-067448/07.
XX	
DR	N-PSDB; AEF05327.
XX	
PT	New polynucleotide encoding Protective Antigen, useful for inducing an
XX	immune response and for protecting against infection by Bacillus
XX	anthracis.
XX	
PS	Claim 4; SEQ ID NO 5; 110pp; English.
XX	

The new invention relates to sequences encoding recombinant *Bacillus anthracis* protective antigen (rPA). Specifically described is a nucleic acid sequence (SEQ ID NO: 1), encoding recombinant rPA, or a fragment of the nucleic acid. Also described is an expression vector comprising the new polynucleotide; a host cell comprising the expression vector; a method for producing rPA or its fragment; a polypeptide produced by the method of (3) or which is encoded by the new polynucleotide; a kit comprising one or more of the polynucleotide, expression vector, host cell, or polypeptide; an antigenic composition comprising a polypeptide of (4); and a method of inducing an immune response against infection by *B. anthracis*. The polynucleotide sequence further comprises a polynucleotide encoding a secretion sequence, e.g., ompA, ompT, ompF, pelB, phoA, lamb, beta lactamase, *Staphylococcus aureus* protein A, *Bacillus subtilis* endoglucanase, murine RNase, human growth hormone, enterotoxin ST-II, LT-A or LT-B, or cpg 2 (cpg). The secretion sequence is for periplasmic or extracellular translocation of the expressed polypeptide from a bacterial host cell cytoplasm. The secretion sequence is a cpg 2 (cpg) leader sequence encoded by SEQ ID NO: 3. The polypeptide is useful for manufacturing a medicament for inducing an immune response against infection by *B. anthracis*, and for protecting against infection by *B. anthracis*. The present sequence is recombinant Protective Antigen, rPA, protein sequence.

Sequence 735 AA;

Query Match	96.4%	Score 3634	DB 10	Length 735
Best Local Similarity	97.2%	Pred. NO. 4.6e-218		
Matches 717, Conservative	4	Mismatches 11	Indels 6	Gaps 3

QY	1	EYKQENRRLNEBSSSOGLGYTFEDLNFQAAMVYTSSTTGGDLSFSSLELNPSPNOVF	60
Db	1	EYKQENRRLNEBSSSOGLGYTFEDLNFQAAMVYTSSTTGGDLSFSSLELNPSPNOVF	60
QY	61	QSAIWSGFLKYKXSGDEYFATSDADNHVTVMVDDOEYINKASNSNKRILEKGRLOYKIQY	120
Db	61	QSAIWSGFLKYKXSGDEYFATSDADNHVTVMVDDOEYINKASNSNKRILEKGRLOYKIQY	120
QY	121	QRENTEKGLDFKLWYTDSONKCEVISSDNLQPELKQKSSNSRKKRSTISAGPTVPDRN	180
Db	121	QRENTEKGLDFKLWYTDSONKCEVISSDNLQPELKQKSSNSRKKRSTISAGPTVPDRN	180
QY	181	DGIPSLFVEGTVVQKXKRTFLSMWINSIHKKGLTKYKSPPEKMSSTDSPDPEKXT	240
Db	181	DGIPSLFVEGTVVQKXKRTFLSMWINSIHKKGLTKYKSPPEKMSSTDSPDPEKXT	240
QY	241	GRIDKNSVPEARHPLVAAPYPIVHVMMENILSKNEDOSTONTDSQTRITISKYSTSRDAN	300
Db	241	GRIDKNSVPEARHPLVAAPYPIVHVMMENILSKNEDOSTONTDSQTRITISKYSTSRDAN	300
QY	301	TWGVISISAGYONGF--TGNITTSAGFSNSNSSTVALDHSLSLAGERTVAETMGLNTAOT	357
Db	301	TWGVISISAGYONGF--TGNITTSAGFSNSNSSTVALDHSLSLAGERTVAETMGLNTAOT	357
QY	358	ARLNANIHYVNTGTAPINVLPTTSVLVCKNOTLATIKAKENQOLSOILAPNNYPPSKNLA	417
Db	358	ARLNANIHYVNTGTAPINVLPTTSVLVCKNOTLATIKAKENQOLSOILAPNNYPPSKNLA	417
QY	418	PIALNAODPFSSPTITMNYNOFLBLEKTKQRLDPTDQVYGNATYTNFENGVRVDTGSNW	477
Db	418	PIALNAODPFSSPTITMNYNOFLBLEKTKQRLDPTDQVYGNATYTNFENGVRVDTGSNW	477
QY	478	SEVLPJOIETTARILFNGKDLNVERRILAAVNSPBLETTKDDMTLKBALKILAFGENBN	537
Db	478	SEVLPJOIETTARILFNGKDLNVERRILAAVNSPBLETTKDDMTLKBALKILAFGENBN	537
QY	538	GNLQYQGDITFEDPNPDQOTSQNIKNQOLAEIATNITYVLDKIKLANAKNILLIRDKRFH	597
Db	538	GNLQYQGDITFEDPNPDQOTSQNIKNQOLAEIATNITYVLDKIKLANAKNILLIRDKRFH	597
QY	558	YDRNNIANGABESVYKEARREYVINSTGGLNIDDKIRKILISGYIVLEIDEBGLKEVIN	657
Db	558	YDRNNIANGABESVYKEARREYVINSTGGLNIDDKIRKILISGYIVLEIDEBGLKEVIN	657

QY 658 DRYDMLNTSLRQDGKTFIDPKKYNKPLPYISNPYKVVYAVTKENTIIINPSNGDTS 717
DB 658 DRYDMLNTSLRQDGKTFIDPKKYNKPLPYISNPYKVVYAVTKENTIIINPSNGDTS 717
QY 718 TNGIKKILIFSKKGYEIG 735
DB 718 TNGIKKILIFSKKGYEIG 735
RESULT 9
ID AEF05332 standard; protein; 736 AA.
AC AEF05332;
XX AEF05332;
XX 23-FEB-2006 (first entry)
DE Wild type Protective Antigen, rPA, SEQ ID NO:6.
XX protective antigen; PA; antigen; Bacillus anthracis infection;
KM antibacterial; infection.
XX Bacillus anthracis.
XX WO2005123764-A1.
XX 29-DEC-2005.
XX 16-JUN-2005; 2005WO-GB002360.
XX 16-JUN-2004; 2004GB-00013475.
XX 16-JUN-2004; 2004US-0579687P.
XX 15-JUN-2005; 2005US-00153865.
XX (HEALTH PROTECTION AGENCY.
PA (DYNP-) DYNPOT VACCINE CO LLC.
PI Brehm J, Mcentee I, Vincent P, Allison N, Brehm R, Jack G;
PI Herbert M, Solow BT, Arroyo J, Lapevich RK;
XX WPI; 2006-067448/07.
XX New polynucleotide encoding Protective Antigen, useful for inducing an
PT immune response and for protecting against infection by Bacillus
PT anthracis.
PS Disclosure; SEQ ID NO 6; 110pp; English.
XX The new invention relates to sequences encoding recombinant Bacillus
CC anthracis protective antigen (rPA). Specifically described is a nucleic
CC acid sequence (SEQ ID NO: 1), encoding recombinant rPA, or a fragment of
CC the nucleic acid. Also described is an expression vector comprising the
CC new polynucleotide; a host cell comprising the expression vector; a
CC method for producing rPA or its fragment; a polypeptide produced by the
CC method of (3) or which is encoded by the new polynucleotide; a kit
CC comprising one or more of the polynucleotide, expression vector, host
CC cell, or polypeptide; an antigenic composition comprising a polypeptide
CC of (4); and a method of inducing an immune response against infection by
CC B. anthracis. The polynucleotide sequence further comprises a
CC polynucleotide encoding a secretion sequence, e.g. ompA, ompT, ompF,
CC pelB, phoA, lamb, beta lactamase, Staphylococcus aureus protein A,
CC Bacillus subtilis endoglucanase, murine RNase, human growth hormone,
CC enterotoxin ST-II, LT-A or LT-B, or cpg 2 (cpg). The secretion sequence
CC is for periplasmic or extracellular translocation of the expressed
CC polypeptide from a bacterial host cell cytoplasm. The secretion sequence
CC is a cpg 2 (cpg) leader sequence encoded by SEQ ID NO: 3. The polypeptide
CC is useful for manufacturing a medicament for inducing an immune response
CC against infection by B. anthracis, and for protecting against infection
CC by B. anthracis. The present sequence is the wild type Protective
CC Antigen, rPA, protein sequence.
XX Sequence 736 AA;
SQ

Query Match 96.4%; Score 3634; DB 10; Length 736;
Best Local Similarity 97.2%; Pred. No. 4,66-218;
Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;
QY 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVTSTTGDLSIPSELENIPSENOYF 60
DB 2 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVTSTTGDLSIPSELENIPSENOYF 61
QY 61 QSAIWSGFIKYKSDDEYTFATSDNHYTMVDDDEVINKASNKIRLEKRLVOIKYQ 120
DB 62 QSAIWSGFIKYKSDDEYTFATSDNHYTMVDDDEVINKASNKIRLEKRLVOIKYQ 121
QY 121 QRENPTKGLDFKLYMTDSQNKKEVISSDNQLPELKQKSNSRKKSTSGAPTPPBDN 180
DB 122 QRENPTKGLDFKLYMTDSQNKKEVISSDNQLPELKQKSNSRKKSTSGAPTPPBDN 181
QY 181 DGIPLSLEVEGYVDVKNKRTFSLSPWISNIHEKGLTKYKSSPERKWTASDPYDPEY 240
DB 182 DGIPLSLEVEGYVDVKNKRTFSLSPWISNIHEKGLTKYKSSPERKWTASDPYDPEY 241
QY 241 GRIDKNVSPKARHPLVAAVPIVHDMENIILSKNEDOSTONTDSQTRTISKNTSRDAN 300
DB 242 GRIDKNVSPKARHPLVAAVPIVHDMENIILSKNEDOSTONTDSQTRTISKNTSR-TH 300
QY 301 TVGVISAGYONGF---TGNITTSAGEFSNSNSTVAIDHSLSLAGERTWAEIMGINTADT 357
DB 301 TSEVHGNAEYHVASFPDYGSV--SAGEFSNSNSTVAIDHSLSLAGERTWAEIMGINTADT 358
QY 358 ARLNANIRYVNTGAPIVNVLPPTSLVLGKNQTLATTIKAKENQLSQIAPANNYPSKULA 417
DB 359 ARLNANIRYVNTGAPIVNVLPPTSLVLGKNQTLATTIKAKENQLSQIAPANNYPSKULA 418
QY 418 PIALNAODPESSTPTTMVNOFLELEKTKOURLTDDVYGNIAIATYNPENGVRVDTSNM 477
DB 419 PIALNAODPESSTPTTMVNOFLELEKTKOURLTDDVYGNIAIATYNPENGVRVDTSNM 478
QY 478 SEVLPOIETTARIIIPNGKDLNVERRIAANVPSDPLETTPMDTLKEALKIAGFNEPN 537
DB 479 SEVLPOIETTARIIIPNGKDLNVERRIAANVPSDPLETTPMDTLKEALKIAGFNEPN 538
QY 538 GNLOYGKDITEFPNPDQOTSQNIKNQALAEANITVTVDKIKLANAKNIIIRDGRFH 597
DB 539 GNLOYGKDITEFPNPDQOTSQNIKNQALAEANITVTVDKIKLANAKNIIIRDGRFH 598
QY 598 YDRNNIANGADESVYKKAHREVINSSTEGLLNTDKDIRKILSGIYVIEDTEGLKXYIN 657
DB 599 YDRNNIANGADESVYKKAHREVINSSTEGLLNTDKDIRKILSGIYVIEDTEGLKXYIN 658
QY 658 DRYDMLNTSLRQDGKTFIDPKKYNKPLPYISNPYKVVYAVTKENTIIINPSNGDTS 717
DB 659 DRYDMLNTSLRQDGKTFIDPKKYNKPLPYISNPYKVVYAVTKENTIIINPSNGDTS 718
QY 718 TNGIKKILIFSKKGYEIG 735
DB 719 TNGIKKILIFSKKGYEIG 736
RESULT 10
ID AAE35717
AC AAE35717 standard; protein; 764 AA.
XX AAE35717;
XX 17-JUN-2003 (first entry)
DE Bacillus anthracis protective antigen.
XX Apoptosis; therapy; inflammatory mediator; intracellular trafficking;
KM infection; prion disease; Alzheimer' disease; hypersecretion disorder;
KM muscle spasm; COPD; bronchitis; chronic obstructive pulmonary disease;
KM torticollis; blepharospasm; asthma; protective antigen.
OS Bacillus anthracis.

PN WO200296467-A2.

PD 05-DEC-2002.

XX 21-MAY-2002; 2002WO-GB002384.

PF 24-MAY-2001; 2001GB-00012687.

PR (MICR-) MICROBIOLOGICAL RES AUTHORITY.

PA Sutton JM, Shone CC;

PI WPI: 2003-167247/16.

XX Conjugate for modulating cell survival and cell growth, modulating

PT release of inflammatory mediator from cells, comprises injected bacterial

PT effector protein and a carrier that targets the protein to target cell.

PS Example 12; Page 122-125; 130pp; English.

XX The invention relates to a conjugate comprising an injected bacterial

XX effector protein and a carrier that targets the effector protein to a

CC target cell. Pharmaceutical composition of the invention is useful for a

CC treatment selected from promoting or inhibiting survival of cells;

CC preventing and reversing damage to cells; killing cells; promoting or

CC inhibiting the growth of cells; apoptosis, release of an inflammatory

CC mediator from cells; division of cells and treating intracellular

CC infection and regulating nitric oxide release from cells. The invention

CC is useful in the manufacture of a medicament for treating a neuronal

CC cell, for intracellular infection, for interfering with intracellular

CC trafficking, for modulating expression of cell-surface markers and for

CC inhibiting secretion from cells. The invention is also useful for

CC treating Prion disease, Alzheimer' disease and wide range of disorders

CC including muscle spasms such as blepharospasm, torticollis and

CC hypersecretion disorders such as chronic obstructive pulmonary disease

CC (COPD), bronchitis and asthma. The present sequence is Bacillus anthracis

CC protective antigen. This sequence is used in the exemplification of the

XX invention

XX

SQ Sequence 764 AA;

Query Match 96.4%; Score 3634; DB 6; Length 764;

Best Local Similarity 97.2%; Pred. No. 4, 9e-218;

Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3

QY 1 EVKQENRLNSESSESSOGGLGYPSFDLNFQAPMYVTSSTWTGDLSPSELXNIPSENQYF 60

DB 30 EVKQENRLNSESSESSOGGLGYPSFDLNFQAPMYVTSSTWTGDLSPSELXNIPSENQYF 89

QY 61 QSAIWGSGFIKYKKSDEYTFATSDAHYVMVVDQEVINKASNNKIRLEKGLYQIKIY 120

DB 90 QSAIWGSGFIKYKKSDEYTFATSDAHYVMVVDQEVINKASNNKIRLEKGLYQIKIY 149

QY 121 QRENPTXKGLDPKLYMTDSONKKEVISSDNIQLPEIKQSSNSRKKRSTSAQPTVPDRDN 180

DB 150 QRENPTXKGLDPKLYMTDSONKKEVISSDNIQLPEIKQSSNSRKKRSTSAQPTVPDRDN 209

QY 181 DGIPTDLSLEVEGYTVDVKNKRPFLSMWISINHEKKGLTKYKSSPEKMWSTADSPDEKYT 240

DB 210 DGIPTDLSLEVEGYTVDVKNKRPFLSMWISINHEKKGLTKYKSSPEKMWSTADSPDEKYT 269

QY 241 GRIDKNVSPPEARHPVLAAYPIVHVWMENTILSKNEQDOSTQNTDQSTRTSKNYSRDN 300

DB 270 GRIDKNVSPPEARHPVLAAYPIVHVWMENTILSKNEQDOSTQNTDQSTRTSKNYSRDN 328

QY 301 TVGVVISAGYQNGF--TGNITTSAGFSNSNSIYADHSLSLAGERTVAETMGLNTADT 357

DB 329 TSEVHGNALEVAHSPFDIGGSV--SAGFSNSNSIYADHSLSLAGERTVAETMGLNTADT 386

QY 358 ARLANANIVYVGTGTPPIYVLPFTTSLVGLKNOTLATIRAKENQLSQILAPNNYIPSKNLA 417

DB 387 ARLANANIVYVGTGTPPIYVLPFTTSLVGLKNOTLATIRAKENQLSQILAPNNYIPSKNLA 446

	QY	416 PLAINADDPSSPTITMANNOGLELEKTKOLRLTDVGYGNIAIYYENGVRVDTSNM	47
	Dd	447 PIALNADDDPSSPTITMANNOPLELEKTOKLRLTDOVGNIATYYENGRVVDTSNW	506
	QY	478 SEVLPOIEETTRARILFNGKDNLVRRRIAAVNPSDPLETTTPDMTLKEALKIAGFNBN	537
	Dd	507 SBLPLIOGOTTARRILFNGKDNLVERRIAANVPSPLETTFPDMTLKGLKIAGFNEPN	566
	QY	538 GNILOYGDITEFDNFDDQSQNINKOALSNATNYITYUDIKLAANKMIILRDGRFH	597
	Dd	567 GNLOYGODITEFDNFDDQSQNINKOALSNATINYIVYIDKLKLANKMILLRDRGFH	626
	QY	598 YDBNNIVAADSVAEAREVINYSTGGLLLNLDKOIRKLISGIYEIEDTEGAKEYIN	657
	Dd	627 YDBNNIVAADSVAEAREVINSTGCLLNLDKDIRKLISGIYIEEDTEGAKEYIN	686
	QY	658 DRYDMNISSLRKDGKITDFPKCYNDKLPLYISMNPVKVVNAVTKENTIIINPSENGDT	717
	Dd	687 DRDYDNMISSLSAQDKGITDFPKCYNCKLPPLYISMNPVKVVNAVTKENTIINPSENGDT	746
	OY	718 TNGIKKKILLFSKKGYBIG 735 	
	Dd	747 TNGIKKKILLFSKKGYBIG 764 	
	RESULT_11		
	ADI95301		
ID	ADI95301 standard; protein; 764 AA.		
XZ	ADI95301;		
DT	04-NOV-2004 (first entry)		
DE	B anthracis toxin translocating protein -protective antigen precursor PA.		
KW	immune response; overlapping synthetic peptide formulation; OSPF; immunostimulant; virucide; antibacterial; antiparasitic; cytoplastic; vaccine; viral; bacterial; parasitic infection; prion disease; neoplastic; toxin translocating protein protective antigen precursor PA. Bacillus anthracis.		
PX	WO2004002415-A2.		
PN	08-JAN-2004.		
PD	27-JUN-2003; 2003WO-US020322.		
PF	27-JUN-2002; 2002US-039271BP.		
PR	(DAHD) DNA FARMER CANCER INST INC.		
PX	Ruprecht RW, Jiang S; WPI; 2004-082868/08.		
PA			
PX			
PI			
PJ			
PT	Modulating an immune response, useful for treating immune disorders, e.g. viral, bacterial and parasitic infections, prion diseases, or neoplastic diseases, administering to a subject an overlapping synthetic peptide formulation.		
PT			
PS	Claim 13; SEQ ID NO 209; 175pp; English.		
CX	The invention relates to a novel method for modulating an immune responsee comprising administering to a subject an overlapping synthetic peptide formulation (OSPF) which comprises a combination of single chain peptides corresponding to the amino acid sequence of a protein of interest. The method of the invention has immunostimulant, virucide, antibacteriacal, antiparasitic and cycostatic applications and may be useful during vaccine production and for treating immune disorders including viral, bacterial and parasitic infections, prion diseases, neoplasiic diseases, as well as providing protection against toxins. The current sequence is		

CC that of the OSPF-related Bacillus anthracis toxin translocating protein
CC (protective antigen precursor [PA]) of the invention.

XX Sequence 764 AA;

Query Match 96.4%; Score 3634; DB 8; Length 764;

Best Local Similarity 97.2%; Pred. No. 4.9e-218;

Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;

```
QY      1 EVKQENRLINSESSSQGLGYPSDLNFOAPMVVTSSTTGDLSIPSELENIPSENGYF 60
DB      30 EVKQENRLINSESSSQGLGYPSDLNFOAPMVVTSSTTGDLSIPSELENIPSENGYF 89
QY      61 OSATWSGRTKYKKSDEYTPATSADNHVTMWDDQVINKASNSNKTREKGLVQIKIY 120
DB      90 OSATWSGRTKYKKSDEYTPATSADNHVTMWDDQVINKASNSNKTREKGLVQIKIY 149
QY     121 QRENPTKGLDFKLWYTSQNKKEVISSDNLOLPBLKOKSSNSRKRKSTASGPTVPDRDN 180
DB     150 QRENPTKGLDFKLWYTSQNKKEVISSDNLOLPBLKOKSSNSRKRKSTASGPTVPDRDN 209
QY     181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDPEKVT 240
DB     210 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDPEKVT 269
QY     241 GRIDKNVSPBARHPIVAAVPIVHVDMENTILSKNEDOSTQNTDSTRTISKNTSTSRDN 300
DB     270 GRIDKNVSPBARHPIVAAVPIVHVDMENTILSKNEDOSTQNTDSTRTISKNTSTSR 328
QY     301 TVGVASISAGYQNGF--TGNITTSAGFSNSNSTVAIDHSLSLAEERTWAETMGINTADT 357
DB     329 TSEVHGNAEVAASFDPDIGSV--SAGFSNSNSTVAIDHSLSLAEERTWAETMGINTADT 386
QY     358 ARLANINIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIRAKENQLSQIIAPNNYPSKNIA 417
DB     387 ARLANINIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIRAKENQLSQIIAPNNYPSKNIA 446
QY     418 PIALNAODPFSTPTTMMYNOPLLEKTKQKLPDIDVOYGNIATYNFENGRRVDTGSNW 477
DB     447 PIALNAODPFSTPTTMMYNOPLLEKTKQKLPDIDVOYGNIATYNFENGRRVDTGSNW 506
QY     478 SEVLPQIDETTRARIIFNGKDLNVERRIAANVSPDLETTEKDMTLKELAKIAFGFNEPN 537
DB     507 SEVLPQIDETTRARIIFNGKDLNVERRIAANVSPDLETTEKDMTLKELAKIAFGFNEPN 566
QY     538 GNLQYQSKDITFEDFNPDQOTSQNIKNQLABELANINITYVLDKIKLANAMNILLRDKRFH 597
DB     567 GNLQYQSKDITFEDFNPDQOTSQNIKNQLABELANINITYVLDKIKLANAMNILLRDKRFH 626
QY     598 YDRNNIATGADSVYKEAREVYINSTGBLLINDIKRILISGTVIEBDETGKKEYVN 657
DB     627 YDRNNIATGADSVYKEAREVYINSTGBLLINDIKRILISGTVIEBDETGKKEYVN 686
QY     658 DRYDMLNISTSLKQDKFTIDFKKNDKLPYISNPNYKVVAVYAKENTIIINPSNGDTS 717
DB     687 DRYDMLNISTSLKQDKFTIDFKKNDKLPYISNPNYKVVAVYAKENTIIINPSNGDTS 746
QY     718 TNGIKKILIFSKKGYEIG 735
DB     747 TNGIKKILIFSKKGYEIG 764
RESULT 12
ID      ADMW29248 standard; protein; 764 AA.
XX      ADMW29248;
XX      24-MAR-2005 (first entry)
XX      B. anthracis protective antigen, PA.
XX      Protective antigen; bacillus anthracis infection; vaccine; infection;
```

KW antibacterial.

OS Bacillus anthracis.

PN WO200500884-A1.

PD 06-JAN-2005.

PP 04-JUN-2004; 2004WO-US017736.

PR 05-JUN-2003; 2003US-0476598P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Schneerson R, Lepja S, Robbina JB, Shiloach J, Kubler-Kielb J;

PI Liu D, Majdady F;

DR WPI; 2005-066555/07.

XX N-PSDB; ADMW29247.

PT Immunogenic conjugate, useful for inhibiting or treating infection and/or
PT disease caused by Bacillus anthracis or other bacilli, comprises Bacillus
PT capsular poly-gamma-glutamic acid polypeptide covalently linked to
PT carrier.

PS Disclosure; SEQ ID NO 3; 66pp; English.

CC The invention relates to an immunogenic conjugate comprising a Bacillus
CC capsular poly-gamma-glutamic acid (gammaPGA) polypeptide covalently
CC linked to a carrier, where the conjugate elicits an immune response in a
CC subject. Also included are a composition comprising the conjugate and a
CC pharmaceutical acceptable carrier (for use in eliciting an immune
CC response against Bacillus antigenic epitope, preferably Bacillus
CC anthracis antigenic epitope in a subject) and an isolated antibody that
CC binds to the Bacillus capsular gammaPGA polypeptide (that recognizes
CC antigenic epitopes on both the Bacillus capsular gammaPGA polypeptide,
CC or that binds Bacillus anthracis capsular gammaPGA). The carrier may
CC comprise recombinant B. anthracis protective antigen. The composition and
CC conjugate are useful for eliciting an immune response against a Bacillus
CC antigenic epitope in a subject and for inhibiting or treating infection
CC and/or disease caused by B. anthracis or other bacilli. The immune
CC response comprises opsonophagocytic activity. The composition is useful
CC for active immunization for preventing B. anthracis infections and for
CC preparation of immune antibodies. The present sequence represents
CC Bacillus anthracis protective antigen, PA.

XX Sequence 764 AA;

Query Match 96.4%; Score 3634; DB 9; Length 764;

Best Local Similarity 97.2%; Pred. No. 4.9e-218;

Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;

```
QY      1 EVKQENRLINSESSSQGLGYPSDLNFOAPMVVTSSTTGDLSIPSELENIPSENGYF 60
DB      30 EVKQENRLINSESSSQGLGYPSDLNFOAPMVVTSSTTGDLSIPSELENIPSENGYF 89
QY      61 OSATWSGRTKYKKSDEYTPATSADNHVTMWDDQVINKASNSNKTREKGLVQIKIY 120
DB      90 OSATWSGRTKYKKSDEYTPATSADNHVTMWDDQVINKASNSNKTREKGLVQIKIY 149
QY     121 QRENPTKGLDFKLWYTSQNKKEVISSDNLOLPBLKOKSSNSRKRKSTASGPTVPDRDN 180
DB     150 QRENPTKGLDFKLWYTSQNKKEVISSDNLOLPBLKOKSSNSRKRKSTASGPTVPDRDN 209
QY     181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDPEKVT 240
DB     210 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDPEKVT 269
QY     241 GRIDKNVSPBARHPIVAAVPIVHVDMENTILSKNEDOSTQNTDSTRTISKNTSTSRDN 300
DB     270 GRIDKNVSPBARHPIVAAVPIVHVDMENTILSKNEDOSTQNTDSTRTISKNTSTSR 328
QY     301 TVGVASISAGYQNGF--TGNITTSAGFSNSNSTVAIDHSLSLAEERTWAETMGINTADT 357
```


Db 329 TSEVHGMAFHAHAFPDGSSV--SAGSNNSSSTVAIDHSLAGERTAFMTGINTADT 386
 QY 358 ARLANANIRYVNTGAPLYNVLPPTSLVAGKNOQLATIKAKENOLSOILAPNNYPSKNLA 417
 Db 387 ARLANANIRYVNTGAPLYNVLPPTSLVAGKNOQLATIKAKENOLSOILAPNNYPSKNLA 446
 QY 418 PILANNOODSSSPITMNNVNOELETKQLRLDPTQVGNITATYNNENRNVVDGTSNW 477
 Db 447 PILANNOODSSSPITMNNVNOELETKQLRLDPTQVGNITATYNNENRNVVDGTSNW 506
 QY 478 SEVLPOIOETTAIRIFNGKDLNVERRIAANVNSDPLETTPDMTLKALIKAFGEFEPN 537
 Db 507 SEVLPOIOETTAIRIFNGKDLNVERRIAANVNSDPLETTPDMTLKALIKAFGEFEPN 566
 QY 538 GNLQYQKDIETEPFNPDOQTSQNIKNQALNATNITVYLDKIKLNAXNNILIRDRFH 597
 Db 567 GNLQYQKDIETEPFNPDOQTSQNIKNQALNATNITVYLDKIKLNAXNNILIRDRFH 626
 QY 598 YDRNNINAVGADSEVYKAAHREVNINSSTEGLLNIDKDIRKILSGYIYEIDTGLKEVIN 657
 Db 627 YDRNNINAVGADSEVYKAAHREVNINSSTEGLLNIDKDIRKILSGYIYEIDTGLKEVIN 686
 QY 658 DRYDMLNISLRDQKTFIDFKKYNDKPLIYISNPYKVVAVVTKENTINSENGDTS 717
 Db 687 DRYDMLNISLRDQKTFIDFKKYNDKPLIYISNPYKVVAVVTKENTINSENGDTS 746
 QY 718 TNGIKKILIFSCKGYEIG 735
 Db 747 TNGIKKILIFSCKGYEIG 764
 RESULT 13
 AED17612
 ID AED17612 standard; protein; 764 AA.
 AC AED17612;
 XX 01-DEC-2005 (first entry)
 DT Anthrax toxin protective antigen PrAg.
 DE Anthrax toxin protective antigen PrAg.
 XX Anthrax toxin; protein engineering; cancer; tumor; neoplasm; cytostatic;
 KM viral infection; virulence; anti-HIV; autoimmune disease; immune disorder;
 KM antitubercular; antidiabetic; antidiabetic; muscular-gen.;
 KM neuroprotective; antiinflammatory; dermatological; antithyroid; anabolic;
 KM immunosuppressive; hypertensive.
 XX Bacillus anthracis.
 OS Bacillus anthracis.
 XX Key
 FH Location/Qualifiers
 FT 1. 29
 FT /label= Signal_peptide
 FT 30..764
 FT /label= Mature_PrAg
 FT Cleavage-site
 FT 193..196
 FT /note= "Furin cleavage site"
 PN MO2005090393-A2.
 XX 29-SEP-2005.
 PD 09-FEB-2005; 2005WO-US004216.
 XX 09-FEB-2004; 2004US-0543417P.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Lepria SH, Liu S, Bugge TH;
 PI WPI; 2005-649596/66.
 DR N-PSDB; AED17611.
 XX

PT Composition comprising a first effector component of a multimeric
 PT bacterial protein toxin, useful for treating cancer including melanoma or
 PT breast cancer, viral infection, or autoimmune disease e.g. Rheumatoid
 PT arthritis.
 PS Disclosure; SEQ ID NO 12; 83pp; English.
 XX
 CC The present invention provides modified bacterial toxins and methods of
 CC using them for targeting specific cell populations. A claimed composition
 CC comprises a first effector component of a multimeric bacterial protein
 CC toxin. The first effector component comprises at least a first monomer
 CC and a different second monomer which form a heterodimer, where the first
 CC and second monomers are each modified by at least 2 of the following
 CC methods: substitution of a native cell-recognition domain for a non-
 CC native cell-recognition domain; substitution of a native proteolytic
 CC activation site for a non-native proteolytic activation site;
 CC modification of the first monomer to generate a first modified monomer,
 CC where the first modified monomer can pair only with the second monomer;
 CC modification of the first monomer and the second monomer, where a second
 CC effector component can bind only at a site formed by the interaction of
 CC the first monomer and the second monomer molecule; or a combination of
 CC these. The first effector component may form a multimeric bacterial
 CC protein toxin component comprising at least 5, 6 or 7 monomers. The
 CC second effector component is selected from anthrax lethal factor (Lpf),
 CC anthrax edema factor, amino acids 1-254 of Lpf or pfs9. The bacterial
 CC protein toxin is selected from anthrax toxin, cholera toxin, Shiga toxin,
 CC tetanospasmin toxin alpha or pertussis toxin. In an exemplary
 CC embodiment, the invention provides modified anthrax protective antigens
 CC (PrAg) that are dependent on intramolecular complementation for toxicity.
 CC These comprise at least one PrAg monomer in which the native furin
 CC cleavage site is replaced with a cleavage site for matrix
 CC metalloproteinase (MMP) and at least one PrAg monomer in which the native
 CC furin cleavage site is replaced with a cleavage site for a plasminogen
 CC activator. The PrAg monomers are further modified by mutation of the
 CC native LP binding site such that at least 2 modified PrAg monomers are
 CC required to bind LP. These modified PrAg specifically target cells
 CC expressing both MMP and plasminogen activators and have reduced toxicity
 CC relative to unmodified PrAg. The claimed composition is used in a claimed
 CC method of treating a disease. The disease is: cancer, such as a
 CC carcinoma, sarcoma, lymphoma or leukemia, especially melanoma, colon
 CC cancer, breast cancer, bladder cancer, thyroid cancer, liver cancer,
 CC pleural cancer, lung cancer, ovarian cancer, pancreatic cancer, head and
 CC neck cancer, kidney cancer, multiple myeloma, stomach cancer, brain
 CC cancer, Hodgkin's lymphoma or non-Hodgkin's lymphoma, where the cancer
 CC cell expresses at least 2 proteolytic enzymes; a viral infection such as
 CC a human immunodeficiency virus (HIV) infection, a cytomegalovirus
 CC infection, a human papillomavirus infection, a hepatitis B virus
 CC infection, a hepatitis C virus infection, a herpes simplex virus
 CC infection or a herpes varicella zoster virus infection, or an autoimmune
 CC disease, such as rheumatoid arthritis, diabetes mellitus, myasthenia
 CC gravis, systemic lupus erythematosus, Grave's disease, or Addison's
 CC disease. Targeting a cell comprises contacting the cell with the
 CC composition. Preferably, the cell is in a mammal. The cell is killed by
 CC the contacting step. The present sequence is the protein sequence of wild
 CC -type PrAg.
 XX
 XX
 XX Sequence 764 AA;
 QY
 Query Match 96.4%; Score 3634; DB 9; Length 764;
 Best Local Similarity 97.2%; Pred. No. 4.9e-218;
 Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;
 QY 1 EVKQENRLNSESSESSGGLAGYFSDLPNFOAPMVVTSSTTGDLSPSSLENIENPSENQYF 60
 Db 30 EVKQENRLNSESSESSGGLAGYFSDLPNFOAPMVVTSSTTGDLSPSSLENIENPSENQYF 89
 QY 61 QSAIWSGFIKVKKSDBEYTPATSDAHNVMTWVDQVYINKASNNKRLREKGLYQIKIY 120
 Db 90 QSAIWSGFIKVKKSDBEYTPATSDAHNVMTWVDQVYINKASNNKRLREKGLYQIKIY 149
 QY 121 QRENPTKGLADPKLWTDSONKKEVYSSDNLQPLKQKSSNRKRSSTASPTVYDRDN 180
 Db 150 QRENPTKGLADPKLWTDSONKKEVYSSDNLQPLKQKSSNRKRSSTASPTVYDRDN 209

```
QY 181 DGIPDSLEVEGYTVDVKKRTFLSPWISNIHEKGLTKYKSPPEKSTASDPYSDFEKT 240
Db 210 DGIPDSLEVEGYTVDVKKRTFLSPWISNIHEKGLTKYKSPPEKSTASDPYSDFEKT 269
QY 241 GRIDKNVSPPEARHPVAAVPIVHVDMENIILSKNEDOSTQNTDSTRTISKTSTSRDAN 300
Db 270 GRIDKNVSPPEARHPVAAVPIVHVDMENIILSKNEDOSTQNTDSTRTISKTSTSR-TH 328
QY 301 TVGYSISAGYONGF--TGNITTSAGFSNSNSTVAIDHSLSLAGERTWAETMGINTADT 357
Db 329 TSEVHGNAEVAASFDDIGGSV--SAGFSNSNSTVAIDHSLSLAGERTWAETMGINTADT 386
QY 358 ARLANIRYVNTGTAPVYVLPVPTSLVIGKQNTLATIRAKENQLSQILAPNNYPSKNLA 417
Db 387 ARLANIRYVNTGTAPVYVLPVPTSLVIGKQNTLATIRAKENQLSQILAPNNYPSKNLA 446
QY 418 PIATNAQDDFSTPTTMMYNNOFLAEKTKQLRLDQVYGNATYVNFENGRRVVDTSNM 477
Db 447 PIATNAQDDFSTPTTMMYNNOFLAEKTKQLRLDQVYGNATYVNFENGRRVVDTSNM 506
QY 478 SEVLPQIETTARIIFNGKDLNVERRIAANVSPDLETTKPDWTLKEALKIAPGNEPN 537
Db 507 SEVLPQIETTARIIFNGKDLNVERRIAANVSPDLETTKPDWTLKEALKIAPGNEPN 566
QY 538 GNLQYQGDITEFDENPDQOQSONIKNOLAEINATYVLDKILNKNMILIDKXPH 597
Db 567 GNLQYQGDITEFDENPDQOQSONIKNOLAEINATYVLDKILNKNMILIDKXPH 626
QY 598 YDRNNIAAGADESVYKEAHREVINSSTEGILLINDIKIRKILSGYIVIEIDTEGLKEVIN 657
Db 627 YDRNNIAAGADESVYKEAHREVINSSTEGILLINDIKIRKILSGYIVIEIDTEGLKEVIN 686
QY 658 DRYDMNLNTSSLRQDKTPIDFKKYNDKLPYISNPYKAVNTAVTKENTINPSSNGDTS 717
Db 687 DRYDMNLNTSSLRQDKTPIDFKKYNDKLPYISNPYKAVNTAVTKENTINPSSNGDTS 746
QY 718 TNGIKKILIFSKGYEIG 735
Db 747 TNGIKKILIFSKGYEIG 764

RESULT 14
AED25934
ID AED25934 standard; protein; 764 AA.
AC AED25934;
DT 01-DEC-2005 (first entry)
XX XX
Db Bacillus anthracis Paga protein.
XX XX
Kw DNA repair; genetic engineering; immune stimulation; immunity;
Kw bacterial infection; infection; antibacterial; Paga; protective antigen.
XX XX
Os Bacillus anthracis.
PN WO2005092372-A2.
PD 06-OCT-2005.
PP 02-FEB-2005; 2005MO-US002987.
XX XX
KX 06-FEB-2004; 2004US-00773618.
PR 30-JUN-2004; 2004US-00863599.
PR 30-JUN-2004; 2004US-0584886P.
PR 23-JUL-2004; 2004MO-US023881.
PR 05-AUG-2004; 2004US-05995222P.
XX XX
PA (CERU-) CERUS CORP.
PA (REGC) UNIT CALIFORNIA.
RI Dubensky TW, Portnoy DA, Calendar RL, Cook DN, Hearst JB;
```

```
XX MPI: 2005-703101/72.
DR N-BSDb; AED25935.
DR Novel isolated, asporogenic Bacillus anthracis strain being attenuated
PT for nucleic acid repair and comprising nucleic acid encoding protective
PT antigen, useful for inducing immune response against B.anthraxis.
XX
PS Disclosure; SEQ ID NO 18; 133pp; English.
XX
CC The invention relates to a modified asporogenic Bacillus anthracis strain
CC (1) that is attenuated for nucleic acid repair e.g. nucleotide excision
CC repair (NER), where the strain is: (a) defective with respect to at least
CC one DNA repair enzyme chosen from UvrA, UvrB, UvrC and RecA, (b)
CC comprising a heterologous expression cassette comprising a nucleic acid
CC encoding a protective antigen, which is operably linked to an inducible
CC promoter, (c) a temperature sensitive recA mutant comprising a recA gene
CC comprising a mutation analogous to the V246M mutation of the recA44
CC temperature sensitive recA mutant of Escherichia coli or recA44(ts)
CC allele of E. coli, (d) a repressible recA mutant, or (e) defective with
CC respect to Stage II sporulation protein B (SpoIIb). Also described are:
CC (1) an isolated, asporogenic B. anthracis bacterium that is attenuated
CC for nucleic acid repair, where the nucleic acid of the B. anthracis
CC bacterium has been modified with a nucleic acid targeting compound that
CC reacts directly with the nucleic acid so that the bacterium is attenuated
CC for proliferation; (2) an isolated, sporulation-deficient B. anthracis
CC strain that is attenuated for NER; (3) a B. anthracis strain being an
CC inducible recA mutant or comprising a nucleic acid encoding an antigen,
CC where the nucleic acid is operably linked to a heterologous promoter; (4)
CC a vaccine composition comprising any one of the modified B. anthracis
CC strains mentioned above or a bacterium from an asporogenic B. anthracis
CC strain that is attenuated for nucleic acid repair; and (5) the use of the
CC bacterium from an asporogenic B. anthracis strain that is attenuated for
CC nucleic acid repair in the manufacture of a medicament for inducing an
CC immune response in a host to B. anthracis. The modified B. anthracis
CC strains of the invention are useful for manufacturing a medicament for
CC inducing an immune response to B. anthracis in a host. The vaccine of the
CC invention is useful for inducing an immune response to B. anthracis in a
CC host. This sequence represents Bacillus anthracis protective antigen
CC (Paga) protein.
XX
SQ Sequence 764 AA;
QY Query Match 96.4%; Score 3634; DB 9; Length 764;
Db Best Local Similarity 97.2%; Pred. No. 4.9e-218;
Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;
QY 1 EVKQENRLNSESSESSQGLGYPSDLNFOAPMVVTSSTGDLSPSSLENNIPSENGYF 60
Db 30 EVKQENRLNSESSESSQGLGYPSDLNFOAPMVVTSSTGDLSPSSLENNIPSENGYF 89
QY 61 OSATWSGPIKYKSDERTFATSADNHTVMWDDQEVINKANSNKRIRLEKRLVQIKIY 120
Db 90 OSATWSGPIKYKSDERTFATSADNHTVMWDDQEVINKANSNKRIRLEKRLVQIKIY 149
QY 121 QRENPTKGLDPFKLYMTDSQNKKEVITSSDNQLPELKOKSSNSRKRSSTAGPTVPDRDN 180
Db 150 QRENPTKGLDPFKLYMTDSQNKKEVITSSDNQLPELKOKSSNSRKRSSTAGPTVPDRDN 209
QY 181 DGIPDSLEVEGYTVDVKKRTFLSPWISNIHEKGLTKYKSPPEKSTASDPYSDFEKT 240
Db 210 DGIPDSLEVEGYTVDVKKRTFLSPWISNIHEKGLTKYKSPPEKSTASDPYSDFEKT 269
QY 241 GRIDKNVSPPEARHPVAAVPIVHVDMENIILSKNEDOSTQNTDSTRTISKTSTSRDAN 300
Db 270 GRIDKNVSPPEARHPVAAVPIVHVDMENIILSKNEDOSTQNTDSTRTISKTSTSR-TH 328
QY 301 TVGYSISAGYONGF--TGNITTSAGFSNSNSTVAIDHSLSLAGERTWAETMGINTADT 357
Db 329 TSEVHGNAEVAASFDDIGGSV--SAGFSNSNSTVAIDHSLSLAGERTWAETMGINTADT 386
QY 358 ARLANIRYVNTGTAPVYVLPVPTSLVIGKQNTLATIRAKENQLSQILAPNNYPSKNLA 417
```

Db 387 ARLANIRVYNTGAPIVNVLPTTSLVLGKNQTLATIKAKENQSLQIAFNPNYPSKNLA 446
QY 418 PIALNADDFSSPTITMANNQFLEKTKQLRLDPTQVGNIAIYFENGGRVAVDTGSNW 477
Db 447 PIALNADDFSSPTITMANNQFLEKTKQLRLDPTQVGNIAIYFENGGRVAVDTGSNW 506
QY 478 SEVLPOIQTETARIIFNGKQNLVVERRIAIVNPSDPLETTKPMUTLKEALKIAGFNEPN 537
Db 507 SEVLPOIQTETARIIFNGKQNLVVERRIAIVNPSDPLETTKPMUTLKEALKIAGFNEPN 566
QY 538 GNLQYQKQITTEPFNFDOQTSQNIKNQALNAIYVVLKIKLNAKNNILIRDRFRH 597
Db 567 GNLQYQKQITTEPFNFDOQTSQNIKNQALNAIYVVLKIKLNAKNNILIRDRFRH 636
QY 598 YDRNNIAVGADESVEKAREVINSSTEGILLNIDKQIRKILSGYIYEIDTEGLKEVIN 657
Db 627 YDRNNIAVGADESVEKAREVINSSTEGILLNIDKQIRKILSGYIYEIDTEGLKEVIN 686
QY 658 DRYDMNISLRODGTFTDFKRYNDKLPYISNPNKVVAVYKENTINSENGDTS 717
Db 687 DRYDMNISLRODGTFTDFKRYNDKLPYISNPNKVVAVYKENTINSENGDTS 746
QY 718 TNGIKKILIFSCKGYEIG 735
Db 747 TNGIKKILIFSCKGYEIG 764
RESULT 15
ID ADR40461 standard; protein; 735 AA.
AC ADR40461;
XX 18-NOV-2004 (first entry)
DE Bacillus anthracis protective antigen mutant F313C.
XX antibacterial; vaccine; protective antigen; PA; proteolytic degradation;
XX B.anthraxis toxin; B.anthraxis infection; passive immunisation;
XX inhalation anthrax; cutaneous anthrax; gastrointestinal anthrax; mutant;
XX mutcin.
OS Bacillus anthracis.
OS Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 313
FT /note= "Wild type Phe substituted by Cys"
XX US2004171121-A1.
XX 02-SEP-2004.
XX 08-AUG-2003; 2003US-00638006.
XX 09-AUG-2002; 2002US-0402285P.
XX (LEPP/) LEPPIA S H.
XX (ROSO/) ROSOVITZ M J.
XX (HSUS/) HSU S D.
XX LepPIA SH, Rosovitz MJ, Hsu SD,
XX WPI; 2004-625107/60.
XX Novel Bacillus anthracis protective antigen having mutations conferring
XX enhanced resistance to proteolytic degradation compared to wild type
XX anthrax toxin.
XX Claim 12; Page: 30pp; English.
XX The invention describes a recombinant Bacillus anthracis protective

CC antigen (PA), modified to incorporate one or more mutations comprising an
CC amino acid deletion or substitution in a flexible, exposed, or loop
CC segment of the PA protein, where the one or more mutations confers
CC enhanced resistance to proteolytic degradation compared to wild type PA.
CC A pharmaceutical composition (i) comprising PA and physiologically
CC acceptable carrier or a composition (ii) comprising a synthetic construct
CC encoding PA and a carrier for inducing serum antibodies that
CC have neutralising activity for a B. anthracis toxin which involves
CC administering (i) to a mammal (human), sufficient to elicit production of
CC the antibodies. The antibodies protect the mammal against the infection.
CC (i) is useful for vaccinating a human against B. anthracis infection. An
CC anti-PA-antibody containing composition is useful for passively
CC immunising a mammal against the toxic effect of B. anthracis. PA is useful
CC as vaccines to induce serum antibodies which are useful to prevent, treat
CC or reduced the severity of infections caused by B. anthracis, such as
CC inhalation anthrax, cutaneous anthrax and/or gastrointestinal anthrax. PA
CC exhibits enhanced stability to proteolytic degradation that is increased
CC by at least 25% compared to proteolytic stability of wild-type PA under
CC comparable conditions. This is the amino acid sequence of protective
CC antigen (PA) mutant F313C. Note: This sequence does not appear in the
CC specification but has been created by the indexer using information given
CC in the specification.
XX
XX SQ Sequence 735 AA;
Query Match 96.4%; Score 3633; DB 8; Length 735;
Best Local Similarity 97.0%; Pred. No. 5,4e-218;
Matches 713; Conservative 5; Mismatches 17; Indels 0; Gaps 0;
QY 1 EVKQENRLNBSRSSGGLGYFSDLNFOAPVWVTSSTGDSIPSSBLENPSENOYF 60
Db 1 EVKQENRLNBSRSSGGLGYFSDLNFOAPVWVTSSTGDSIPSSBLENPSENOYF 60
QY 61 QSAIWSGFIKVKKSDTYTATSADNHVTWVDDQEVINRASNKNKIRLEKRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDTYTATSADNHVTWVDDQEVINRASNKNKIRLEKRLYQIKIY 120
QY 121 QRENPFKEGLDFLWYTDSONKKEVYSSDNQLPELKQKSSNRKRGTSAGTVPDRDN 180
Db 121 QRENPFKEGLDFLWYTDSONKKEVYSSDNQLPELKQKSSNRKRGTSAGTVPDRDN 180
QY 181 DGIPLDSLEVGYYVDVKNKRTPLSPWISNHEKKGLTKYKSSPEKSTASDPSPDEKVT 240
Db 181 DGIPLDSLEVGYYVDVKNKRTPLSPWISNHEKKGLTKYKSSPEKSTASDPSPDEKVT 240
QY 241 GRIDKNVSPRARHPVAAPIYVVDMENIILSKNEQOSTQNTDSQRTTISKNTSTSRDAN 300
Db 241 GRIDKNVSPRARHPVAAPIYVVDMENIILSKNEQOSTQNTDSQRTTISKNTSTSRDAN 300
QY 301 TVGVSIISAGYQNGFTGNITTSAGFSNNSSTVAIDHSLSLAGERTVAETWGLNTADTARL 360
Db 301 SEVHGNAEYVASCDFDGSVSAGFSNNSSTVAIDHSLSLAGERTVAETWGLNTADTARL 360
QY 361 NANIRVYNTGAPIVNVLPTTSLVLGKNQTLATIKAKENQSLQIAFNPNYPSKNLAPIA 420
Db 361 NANIRVYNTGAPIVNVLPTTSLVLGKNQTLATIKAKENQSLQIAFNPNYPSKNLAPIA 420
QY 421 LNAQDDFSSPTITMANNQFLEKTKQLRLDPTQVGNIAIYFENGGRVAVDTGSNWSEV 480
Db 421 LNAQDDFSSPTITMANNQFLEKTKQLRLDPTQVGNIAIYFENGGRVAVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKQNLVVERRIAIVNPSDPLETTKPMUTLKEALKIAGFNEPNGL 540
Db 481 LPOIQTETARIIFNGKQNLVVERRIAIVNPSDPLETTKPMUTLKEALKIAGFNEPNGL 540
QY 541 QYQKQITTEPFNFDOQTSQNIKNQALNAIYVVLKIKLNAKNNILIRDRFRH 600
Db 541 QYQKQITTEPFNFDOQTSQNIKNQALNAIYVVLKIKLNAKNNILIRDRFRH 600
QY 601 NNIAGADESVVEKAREVINSSTEGILLNIDKQIRKILSGYIYEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVEKAREVINSSTEGILLNIDKQIRKILSGYIYEIDTEGLKEVINDRY 660

Qy	661	DLNLSSLRQDGKTRIDPKXNDKLPXYISNPNYKVNYAVTKENTIIINPSNGDSTNG	720
Db	661	DLNLSSLRQDGKTRIDPKXNDKLPXYISNPNYKVNYAVTKENTIIINPSNGDSTNG	720
Qy	721	IKKILIFSKGYEIG	735
Db	721	IKKILIFSKGYEIG	735

Search completed: August 11, 2006, 20:20:52
Job time : 212 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2006, 20:17:36 ; Search time 307 Seconds
(without alignment)
2214.614 Million cell updates/sec

Title: US-10-780-250-1
Perfect score: 3770
Sequence: 1 EYKQENRLINESESSSQGL.....TSTNGIKKILIRSKKGYEIG 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

*Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3634	96.4	736	068GS1_BACAN	068GS1 bacillus an
2	3634	96.4	764	1 PAG_BACAN	P13423 bacillus an
3	3634	96.4	764	2 052NR4_BACAN	052NR4 bacillus an
4	2744	72.8	561	2 042B94_BACAN	042B94 bacillus an
5	2140	56.8	760	2 04MR80_BACCE	04MR80 bacillus ce
6	984	26.1	875	2 046221_CLOPE	046221 clostridium
7	966.5	25.6	876	2 032739_CLODI	032739 clostridium
8	964.5	25.6	876	2 09KH41_CLODI	09KH41 clostridium
9	960.5	25.5	879	2 006498_9MOLU	006498 clostridium
10	829	22.0	721	2 086171_CLOBO	086171 clostridium
11	786	20.8	785	2 05JZY7_9BACL	05JZY7 brevibacill
12	776	20.6	871	2 05JZT2_9BACL	05JZT2 brevibacill
13	746	19.8	775	2 0844J8_BACTU	0844J8 bacillus th
14	296.5	7.9	225	1 Y6163_BACAN	P13422 bacillus an
15	232.5	6.2	1305	2 06FNZ8_CANGA	06FNZ8 candida gla
16	214	5.7	1881	2 07ROB6_PLAYO	07ROB6 plasmodium
17	213.5	5.7	1881	2 08RGK2_FUSNN	08RGK2 fusobacteri
18	211.5	5.6	1345	2 04ZIR6_PLABR	04ZIR6 plasmodium
19	209	5.5	4688	2 09P008_UREPA	09P008 ureplasma
20	204.5	5.4	2849	2 081HY4_PLAF7	081HY4 plasmodium
21	203	5.4	1227	2 097K41_CLOAB	097K41 clostridium
22	203	5.4	1965	2 02NFW4_9EURY	02NFW4 methanospa
23	202.5	5.4	1909	2 02NIC8_9EURY	02NIC8 methanospa
24	202.5	5.4	3759	2 07RPU4_PLAYO	07RPU4 plasmodium
25	201.5	5.3	2940	2 081RP9_PLAF7	081RP9 plasmodium
26	201.5	5.3	3119	2 07PES5_FUSNV	07PES5 fusobacteri
27	200.5	5.3	1904	2 08IKS2_PLAF7	08IKS2 plasmodium
28	200	5.3	920	2 07RLG3_PLAYO	07RLG3 plasmodium
29	200	5.3	2529	2 025579_HELPY	025579 helicobacte
30	199.5	5.3	2118	2 07RR57_PLAYO	07RR57 plasmodium
31	199.5	5.3	2331	2 06RMS7_RICTY	06RMS7 rickettsia

32	198.5	5.3	4095	2 07WZ13_HAESO	07WZ13 haemophilus
33	197.5	5.2	1461	2 07RAS6_PLAYO	07RAS6 plasmodium
34	197	5.2	1072	2 04Z640_PLABR	04Z640 plasmodium
35	197	5.2	3248	2 06LFP0_PLAF7	06LFP0 plasmodium
36	196	5.2	2169	2 07RHE7_PLAYO	07RHE7 plasmodium
37	196	5.2	3278	2 04ZHU0_STAHY	04ZHU0 staphylococ
38	196	5.2	10746	2 06GKX3_STNAR	06GKX3 staphylococ
39	195	5.2	2162	2 07RMF9_PLAYO	07RMF9 plasmodium
40	195	5.2	2719	2 07RPU0_PLAYO	07RPU0 plasmodium
41	194.5	5.2	1000	2 07RMA8_PLAYO	07RMA8 plasmodium
42	194.5	5.2	1673	2 06S003_DICDI	06S003 dicystocell
43	194	5.1	1444	2 07PDK0_PLAYO	07PDK0 plasmodium
44	194	5.1	1647	2 07RP46_PLAYO	07RP46 plasmodium
45	193.5	5.1	1231	2 07RT07_PLAYO	07RT07 plasmodium

ALIGNMENTS

RESULT 1					
Q68GS1_BACAN	PRELIMINARY, PRT; 736 AA.				
ID 068GS1; BACAN					
AC 068GS1;					
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.					
DT 11-OCT-2004, sequence version 1.					
DT 07-FEB-2006, entry version 7.					
DE Protective antigen.					
OS Bacillus anthracis.					
OC Bacterioides, Firmicutes; Bacillales; Bacillaceae; Bacillus;					
OC Bacillus cereus group.					
OX NCBI_Taxid=1392;					
OK					
RN [1]					
RP NUCLEOTIDE SEQUENCE.					
RC STRAIN=Stemne;					
RX PubMed=16030177;					
RA Aziz M.A., Sikriwal D., Singh S., Jangula S., Kumar P.A.,					
RA Bhattacharjee R.,					
RT "Transformation of an edible crop with the paxa gene of Bacillus					
RT anthracis."					
RL FASER J. 19:1501-1503 (2005).					
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms					
CC Distributed under the Creative Commons Attribution-NoDerivs license					
CC					
DR EMBL; AY700758; AAT98414.1; -; Genomic DNA.					
DR SMR; Q68GS1; 15-736.					
DR GO; GO:0005576; C:extracellular region; IEA.					
DR GO; GO:0009405; P:pathogenesis; IEA.					
DR InterPro; IPR003896; Anthrax_toxinB.					
DR InterPro; IPR011658; PA14.					
DR InterPro; IPR013145; PA14_rel.					
DR Pfam; PF03495; Binary_toxB; 1.					
DR Pfam; PF07691; PA14; 1.					
DR PRINTS; PR01391; BINARVTOXINB.					
KW Plasmid.					
SQ					
SEQUENCE 736 AA; 82805 MW; D2FF77A38AEE372C CRC64;					
Query Match 96.4%; Score 3634; DB 2; Length 736;					
Best local similarity 97.2%; Pred. No. 1.9e-167;					
Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;					
QY	1 EYKQENRLINESESSSQGLIGYFSDLNQAPMVTSTTTGDLSTSSLENIENPSNOYF 60				
DB	2 EYKQENRLINESESSSQGLIGYFSDLNQAPMVTSTTTGDLSTSSLENIENPSNOYF 61				
QY	61 QSAIWSGFIKKVKSDEYFATSDNHNVTMMVDDQEVYINQASNSNKIRLKKGRLYQIKIQY 120				
DB	62 QSAIWSGFIKKVKSDEYFATSDNHNVTMMVDDQEVYINQASNSNKIRLKKGRLYQIKIQY 121				
QY	121 QRENPEKGLDFKLWYTSQNKKEVSISSDNLQPLBKQSSNSRKKRSTSGAGTVEDRDN 180				
DB	122 QRENPEKGLDFKLWYTSQNKKEVSISSDNLQPLBKQSSNSRKKRSTSGAGTVEDRDN 181				

QY 181 DGIPSLVEGVTTVDVKKRTFTLSPWISNIHEKGLTKYKSPKXSTASDPYSPKXT 240
DB 182 DGIPSLVEGVTTVDVKKRTFTLSPWISNIHEKGLTKYKSPKXSTASDPYSPKXT 241
QY 241 GRIDNGVPEARHPVAAVPIYHVMMENIISKNEDOSTONTDSQTRTSKRTSPRDAN 300
DB 242 GRIDNGVPEARHPVAAVPIYHVMMENIISKNEDOSTONTDSQTRTSKRTSPRDAN 300
QY 301 TVGVISAGYONGF--TGNITTSAGFSNSNSTVAIDHSLSLAGEKRTWAETMGLNTADT 357
DB 301 TSEVHGNAEVAHSFPDIGGSV--SAGFSNSNSTVAIDHSLSLAGEKRTWAETMGLNTADT 358
QY 358 ARLANINITYVTGTPIYVNLPTTSLVIGKQTLATIRAKENQSLQIILAPNNYPSKXLA 417
DB 359 ARLANINITYVTGTPIYVNLPTTSLVIGKQTLATIRAKENQSLQIILAPNNYPSKXLA 418
QY 418 PIALNAODFSSTPIYVNLPTTSLVIGKQTLATIRAKENQSLQIILAPNNYPSKXLA 477
DB 419 PIALNAODFSSTPIYVNLPTTSLVIGKQTLATIRAKENQSLQIILAPNNYPSKXLA 478
QY 478 SEVLPOIETTARIIENGKDLNVERRIAANVSPDLSTTKPDMTLKEALKIAGFNEPN 537
DB 479 SEVLPOIETTARIIENGKDLNVERRIAANVSPDLSTTKPDMTLKEALKIAGFNEPN 538
QY 538 GNLQYQGDITRFDPNPOQTSQNKQLAELANATITVLDKILANAKMILLIDKRFH 597
DB 539 GNLQYQGDITRFDPNPOQTSQNKQLAELANATITVLDKILANAKMILLIDKRFH 598
QY 598 YDRNNIANGADESVYKAEHREYINSTBGLINIDKIRKILSGYIVIEIDTEGKEVIN 657
DB 599 YDRNNIANGADESVYKAEHREYINSTBGLINIDKIRKILSGYIVIEIDTEGKEVIN 658
QY 658 DRYDMLNITSSLQDQKTFIDFKYNDKLPYISNPNYKVVAVTKENTINPSENGDTS 717
DB 659 DRYDMLNITSSLQDQKTFIDFKYNDKLPYISNPNYKVVAVTKENTINPSENGDTS 718
QY 718 TNGIRKILIFSKKGYEIG 735
DB 719 TNGIRKILIFSKKGYEIG 736

RESULT 2
PAG_BACAN STANDARD; PRT; 764 AA.
AC P13423; Q937W2; Q9P5R7; Q9KH69; Q9RDU2;
DT 01-JAN-1990, integrated into UniProtKB/Swiss-Prot.
DT 18-OCT-2001, sequence version 2.
DT 07-MAR-2006, entry version 68.
DE Protective antigen precursor (PA) (PA-83) (PA83) (Anthrax toxin
translocating protein) [Contains: Protective antigen PA-20 (PA20);
Protective antigen PA-63 (PA63)].
GN Name=paga; Synonyms=pag;
Ordered locus names=pxo1-110, BXA0164, GBA0_pxo1_0164;
OS Bacillus anthracis.
OC Plasmid pxo1.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1392;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=89172073; PubMed=3148491; DOI=10.1016/0378-1119(88)90439-8;
RA Welkos S.L., Lowe J.R., Eden-McCuchan F., Vodkin M., Leppla S.H.,
Schmidt J.J.;
RT "Sequence and analysis of the DNA encoding protective antigen of
Bacillus anthracis."
RL Gene 69:287-300(1988).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX STRAIN=28, 33, BA1024, and BA1035;
MEDLINE=99214082; PubMed=10197966;
RA Price L.B., Hugh-Jones M., Jackson P.J., Keim P.;
RT "Genetic diversity in the protective antigen gene of Bacillus

RT anthracis.";
RL J. Bacteriol. 181:2358-2362(1999).
RN [3]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX STRAIN=V770-NP1-R / ATCC 14185;
MEDLINE=20359347; PubMed=10899854;
DOI=10.1128/IAI.68.8.4549-4558.2000;
RA Cohen S., Mendelson I., Altboum Z., Kobiler D., Elhanany E., Bino T.,
Leitner M., Inbar I., Rosenberg H., Gozes Y., Bark R., Fisher M.,
Kotman C., Velan B., Shafterman A.;
RT "Attenuated nontoxicogenic and nonencapsulated recombinant Bacillus
anthracis spore vaccines protect against anthrax."
RL Infect. Immun. 68:4549-4558(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Stierne;
RX MEDLINE=99445483; PubMed=10515943;
Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,
Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,
Martinez Y., Riche D., Svensson R., Jackson P.J.;
RT "Sequence and organization of pxo1, the large Bacillus anthracis
plasmid harboring the anthrax toxin genes."
RL J. Bacteriol. 181:6509-6515(1999).
RN [5]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=Ames / isolate Florida / A2012;
RX MEDLINE=22061436; PubMed=12004073; DOI=10.1126/science.1071837;
Read T.D., Salzberg S.L., Pop M., Shumway M., Umayam L., Jiang L.,
Holtzapple E., Busch J.D., Smith K.L., Schupp J.M., Solomon D.,
Keim P., Frazer C.M.;
RT "Comparative genome sequencing for discovery of novel polymorphisms in
Bacillus anthracis."
RL Science 296:2028-2033(2002).
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Ames ancestor;
RA Ravel J., Raeko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
Frazer C.M.;
RT "Bacillus anthracis comparative genomics."
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 9-751.
RC STRAIN=Carbosap, and Ferrara;
RX MEDLINE=22063192; PubMed=12067380;
DOI=10.1046/j.1365-2672.2002.01650.x;
RA Adome R., Pascual P., La Rosa G., Marinelli C., Muscillo M.,
Faenella A., Francia M., Ciuchini P.;
RT "Sequence analysis of the genes encoding for the major virulence
factors of Bacillus anthracis vaccine strain 'Carbosap'."
RL J. Appl. Microbiol. 93:117-121(2002).
RN [8]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 195-434.
RC STRAIN=PAI;
RA Inoue S., Nouchi A., Tanabayashi K., Yamada A.;
RT "Preparation of a positive control DNA for molecular diagnosis of
Bacillus anthracis."
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN [9]
RP DOMAINS.
RX MEDLINE=91332080; PubMed=1651334;
Singh Y., Klimpel K.R., Quinn C.P., Chaudhary V.K., Leppla S.H.;
RT "The carboxyl-terminal end of protective antigen is required for
receptor binding and anthrax toxin activity."
RL J. Biol. Chem. 266:15493-15497(1991).
RN [10]
RP CHARACTERIZATION.
RC STRAIN=Stierne;
RX MEDLINE=94327640; PubMed=8051159;
Milne J.C., Furlong D., Hanna P.C., Wall J.S., Collier R.J.;
RT "Anthrax protective antigen forms oligomers during intoxication of
mammalian cells."
RL J. Biol. Chem. 269:20607-20612(1994).

RN [11]
RP CHARACTERIZATION. PubMed=11207581;
RX MEDLINE=21129592; PubMed=11207581;
RA DOI=10.1046/j.1462-5822.2000.00052.x;
RA Beaugard K.E., Collier R.J., Swanson J.A.;
RT "Proteolytic activation of receptor-bound anthrax protective antigen
on macrophages promotes its internalization."; Cell. Microbiol. 2:251-258(2000).
RN [12]
RP TOXIN REGULATION.
RC STRAIN=Nybridge;
RX MEDLINE=94131936; PubMed=8300513;
RA Koehler T.M., Dai Z., Kaufman-Yarbray M.;
RT "Regulation of the Bacillus anthracis protective antigen gene: CO2 and
a trans-acting element activate transcription from one of two
promoters."; J. Bacteriol. 176:586-595(1994).
RN [13]
RP FOLDING BY PSRA.
RX MEDLINE=22625622; PubMed=12606539; DOI=10.1074/jbc.M301244200;
RA Williams R.C., Rees M.L., Jacobs M.F., Pragal Z., Thwaitte J.E.,
Baillie L.W., Emerson P.T., Harwood C.R.;
RT "Production of Bacillus anthracis protective antigen is dependent on
the extracellular chaperone, Praa."; J. Biol. Chem. 278:18056-18062(2003).
RN [14]
RP INTERACTION WITH THE ANTHRAX TOXIN RECEPTOR.
RX PubMed=14507921; DOI=10.1074/jbc.M307900200;
RA Bradley K.A., Mogridge J., Jonah G., Rainey G.J.A., Betty S.,
Young J.A.T.;
RT "Binding of anthrax toxin to its receptor is similar to alpha
integrin-ligand interactions."; J. Biol. Chem. 278:49342-49347(2003).
RN [15]
RP MUTAGENESIS OF PHE-342; PHE-343 AND ASP-344.
RX STRAIN=sterne;
MEDLINE=95050722; PubMed=7961869; Sharma M., Leppia S.H.;
RA Singh Y., Kimpel K.R., Aroa N.;
RT "The chymotrypsin-sensitive site, PFD315, in anthrax toxin protective
antigen is required for translocation of lethal factor."; J. Biol. Chem. 269:29039-29046(1994).
RN [16]
RP MUTAGENESIS OF DOMAIN 4 LOOPS.
RX STRAIN=sterne;
MEDLINE=99185012; PubMed=10085028;
RA Varughese M., Teixeira A.V., Liu S., Leppia S.H.;
RT "Identification of a receptor-binding region within domain 4 of the
protective antigen component of anthrax toxin."; Infect. Immun. 67:1860-1865(1999).
RN [17]
RP MUTAGENESIS OF TRP-375; MET-379 AND LEU-381.
RX STRAIN=sterne;
MEDLINE=21092804; PubMed=11178978; DOI=10.1006/bbrc.2001.4320;
RA Batta S., Gupta P., Chauhan V., Singh A., Bhatnagar R.;
RT "Trp 346 and Leu 352 residues in protective antigen are required for
the expression of anthrax lethal toxin activity."; Biochem. Biophys. Res. Commun. 281:186-192(2001).
RN [18]
RP MUTAGENESIS OF PHE-581; PHE-583; ILE-591; LEU-595 AND ILE-603.
RX STRAIN=sterne;
MEDLINE=21348996; PubMed=11554763; DOI=10.1006/bbrc.2001.5613;
RA Ahuja N., Kumar P., Bhatnagar R.;
RT "Hydrophobic residues Phe552, Phe554, Ile562, Leu566, and Ile574 are
required for oligomerization of anthrax protective antigen."; Biochem. Biophys. Res. Commun. 287:542-549(2001).
RN [19]
RP MUTAGENESIS OF PRO-289.
RX STRAIN=sterne;
MEDLINE=21255689; PubMed=11356563;
RA Khanna H., Chopra A.P., Aroa N., Chaudhry A., Singh Y.;
RT "Role of residues constituting the zebrafish strand of domain II in the
biological activity of anthrax protective antigen."; FEMS Microbiol. Lett. 199:27-31(2001).

RN [20]
RP MUTAGENESIS OF GLN-512; ASP-541; LEU-543 AND ARG-621.
RX MEDLINE=21125576; PubMed=11222612;
RX DOI=10.1128/JB.183.6.2111-2116.2001;
RA Mogridge J., Mourez M., Collier R.J.;
RT "Involvement of domain 3 in oligomerization by the protective antigen
molecule of anthrax toxin."; J. Bacteriol. 183:2111-2116(2001).
RN [21]
RP MUTAGENESIS OF LYS-426; ASP-454 AND PHE-456.
RX MEDLINE=21269403; PubMed=11131126; DOI=10.1074/jbc.M008309200;
RA Sellman B.R., Nasri S., Collier R.J.;
RT "Point mutations in anthrax protective antigen that block
translocation."; J. Biol. Chem. 276:8371-8376(2001).
RN [22]
RP MUTAGENESIS OF PRO-213; LEU-216; PHE-231; LEU-232; PRO-234; ILE-236;
RX ILE-239; TRP-255 AND PHE-265.
RX STRAIN=sterne;
MEDLINE=22112896; PubMed=12117959;
RX DOI=10.1128/JAI.70.8.4477-4484.2002;
RA Chauhan V., Bhatnagar R.;
RT "Identification of amino acid residues of anthrax protective antigen
involved in binding with lethal factor."; Infect. Immun. 70:4477-4484(2002).
Query Match 96.4%; Score 3634; DB 1; Length 764;
Best Local Similarity 97.2%; Pred. No. 2e-167;
Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;
QY 1 EYKQENRLNSESSESSQGLGYFFSDLPFOAPVWVTSSTTGDLSTPSSELENTPSSENOYF 60
DB 30 EYKQENRLNSESSESSQGLGYFFSDLPFOAPVWVTSSTTGDLSTPSSELENTPSSENOYF 89
QY 61 QSAIWSGFIKVKKSDTYEATSDNHNVTMVDQEVYNKASNSNKLRLKGLYQIKIYQ 120
DB 90 QSAIWSGFIKVKKSDTYEATSDNHNVTMVDQEVYNKASNSNKLRLKGLYQIKIYQ 149
QY 121 QRENPTKGLDFKLWYTDSONKKEVYSSNQLPELKOKSSNRKRSISAGTVDRDN 180
DB 150 QRENPTKGLDFKLWYTDSONKKEVYSSNQLPELKOKSSNRKRSISAGTVDRDN 209
QY 181 DGIPOSLFVGYTVQVKNKRTPLSPWISNIHEKGLTKKSSPEKSTSDPYDDEKVT 240
DB 210 DGIPOSLFVGYTVQVKNKRTPLSPWISNIHEKGLTKKSSPEKSTSDPYDDEKVT 269
QY 241 GRIDKVSPEARHPVLAAYPIVHVMENIISKNEDOSTQNTDSQRTTISKNTSRDAN 300
DB 270 GRIDKVSPEARHPVLAAYPIVHVMENIISKNEDOSTQNTDSQRTTISKNTSRDAN 328
QY 301 YGVGSISAGYQNGF--TGNITTSAGFSNNSSTVAIDHSLIAGERTVAETWGLNTADT 357
DB 329 TSEVHGNALVHASFPIGDSV--SAGFSNNSSTVAIDHSLIAGERTVAETWGLNTADT 386
QY 358 ARLNARIRVNGTAIVNVLPPTSIVYGNQTLATIKKENQSOILAPNNYPSKNLA 417
DB 387 ARLNARIRVNGTAIVNVLPPTSIVYGNQTLATIKKENQSOILAPNNYPSKNLA 446
QY 418 PIALNAQDPSSPTPIVMANNOFLELEKTKQLRLDTPQVGNATVYFNGRVARVDGSMW 477
DB 447 PIALNAQDPSSPTPIVMANNOFLELEKTKQLRLDTPQVGNATVYFNGRVARVDGSMW 506
QY 478 SEVLPQIGETTAIRIIFNGKOLNIVERRIAAVNSDPLETTKPDMLKEALKIAFGNEEN 537
DB 507 SEVLPQIGETTAIRIIFNGKOLNIVERRIAAVNSDPLETTKPDMLKEALKIAFGNEEN 566
QY 538 GNLQYQGDITFEDRFPDQOTSQNTKNOCLAEINATNIVVLKIKKANAMNIIIRKRF 597
DB 567 GNLQYQGDITFEDRFPDQOTSQNTKNOCLAEINATNIVVLKIKKANAMNIIIRKRF 626
QY 598 YRRNNIAGVADSVVEAREVYNSTEGGLNLINDKIRKILSGYIVLEIDTEGLKEVIN 657
DB 627 YRRNNIAGVADSVVEAREVYNSTEGGLNLINDKIRKILSGYIVLEIDTEGLKEVIN 686

```
QY 658 DRYMLNTSSLRQDKPTIDPKKNDKLPYISNPYKVNYYATKENTIIINPSENGDTS 717
DB 687 DRYMLNTSSLRQDKPTIDPKKNDKLPYISNPYKVNYYATKENTIIINPSENGDTS 746
QY 718 TNGIKKILIFSKKGYEIG 735
DB 747 TNGIKKILIFSKKGYEIG 764

RESULT 3
Q52NH4 BACAN PRELIMINARY; PRT: 764 AA.
AC 052NH4
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DE 07-FEB-2006, entry version 3.
DE Protective antigen.
GN Name: pag;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1392;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AlR;
RA Xu J., Dong D., Chen W.;
RT "Protective antigen gene of Bacillus anthracis strain AlR.";
RL Submitted (Apr-2005) to the EMBL/Genbank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
EMBL: AY97299; AAY15236.1; -; Genomic_DNA.
DB SRR: Q52NH4; 43-764.
GO GO:0005576; C:extracellular region; IEA.
GO GO:0009405; P:pathogenesis; IEA.
DR InterPro: IPR03896; Antirax toxinB.
DR InterPro: IPR013145; PA14_toxI.
DR Pfam: PF03495; BimA_toxI.
DR Pfam: PF07691; PA14_1.
DR PRINTS: PR01391; BINARYTOXINB.
SQ SEQUENCE 764 AA; 85811 MW; 3AB1BFB48FAA03F CRC64;

Query Match 96.4%; Score 3634; DB 2; Length 764;
Beet Local Similarity 97.2%; Pred. No. 2e-167;
Matches 717; Conservative 4; Mismatches 11; Indels 6; Gaps 3;

QY 1 EYKQNRLLNBSSESSOGLAGYFSDLNQAPMVVTSSTTGDLSTIPSELENIPEENQYF 60
DB 30 EYKQNRLLNBSSESSOGLAGYFSDLNQAPMVVTSSTTGDLSTIPSELENIPEENQYF 89
QY 61 QSAIWSGFIKVKSDDEYTPATSDNHNVTWVDQEVINKASNSKIRLEKGLYQIKIY 120
DB 90 QSAIWSGFIKVKSDDEYTPATSDNHNVTWVDQEVINKASNSKIRLEKGLYQIKIY 149
QY 121 QRENTTEKGLDKLWYTSQNKKEVYISSNLQLPFLKQKSSNRKRRKTSAGPTPPDDN 180
DB 150 QRENTTEKGLDKLWYTSQNKKEVYISSNLQLPFLKQKSSNRKRRKTSAGPTPPDDN 209
QY 181 DQIPSLSEVEGTYVVKKRTFLSPWISNIHEKGLTKYKSSPEKWSSTADSYSPFEKYT 240
DB 210 DQIPSLSEVEGTYVVKKRTFLSPWISNIHEKGLTKYKSSPEKWSSTADSYSPFEKYT 269
QY 241 GRIDKNSPEARHPVLAAYPIVHVMENIILSKNEQSTQNTDSQRTISKNTSTRDAN 300
DB 270 GRIDKNSPEARHPVLAAYPIVHVMENIILSKNEQSTQNTDSQRTISKNTSTRDAN 328
QY 301 TVGVYSISAGYQNGF--TGNITTSAGFSNSNSTYALDHSLSLAGERTAAETMGATADT 357
DB 329 TSEVGNNAEVAHSAFDPDGISV--SAGFSNSNSTYALDHSLSLAGERTAAETMGATADT 386
QY 358 ARLNANIRYVNTGTAPIYVNLPTTSLVIGKQNTLATIKAKENQSLQILAANNYYPSKMLA 417
```

```
DB 387 ARLNANIRYVNTGTAPIYVNLPTTSLVIGKQNTLATIKAKENQSLQILAANNYYPSKMLA 446
QY 418 PIALNAQDDFSSPTITWNNYNQFLELEKTKQLRLDTQYGYGAIATYFNENGVRYDYSNW 477
DB 447 PIALNAQDDFSSPTITWNNYNQFLELEKTKQLRLDTQYGYGAIATYFNENGVRYDYSNW 506
QY 478 SEVLPOIQTETARIIFNGKDLNVERRIAANPSDPLETTPEMDTLKALAIARGFENPN 537
DB 507 SEVLPOIQTETARIIFNGKDLNVERRIAANPSDPLETTPEMDTLKALAIARGFENPN 566
QY 538 GNLQYQKDIETEPFENPDQTSQNIKNQLAELNATNIYTVLDKIKLNAKQNILIRDKRFH 597
DB 567 GNLQYQKDIETEPFENPDQTSQNIKNQLAELNATNIYTVLDKIKLNAKQNILIRDKRFH 626
QY 598 YDRNNIAVGADESYYKAEHREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVIN 657
DB 627 YDRNNIAVGADESYYKAEHREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVIN 686
QY 658 DRYMLNTSSLRQDKPTIDPKKNDKLPYISNPYKVNYYATKENTIIINPSENGDTS 717
DB 687 DRYMLNTSSLRQDKPTIDPKKNDKLPYISNPYKVNYYATKENTIIINPSENGDTS 746
QY 718 TNGIKKILIFSKKGYEIG 735
DB 747 TNGIKKILIFSKKGYEIG 764

RESULT 4
Q4ZE94 BACAN PRELIMINARY; PRT: 561 AA.
AC Q4ZE94
DT 07-JUN-2005, integrated into UniProtKB/TrEMBL.
DT 07-JUN-2005, sequence version 1.
DE 07-FEB-2006, entry version 3.
DE Protective antigen (Fragment).
GN Name: pa;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1392;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=34F2;
RA Vahedi F., Moazeni Julia G., Mahmoudi M.;
RT "Humoral immunity in mice induced by vaccination with a plasmid
RT encoding anthrax protective antigen.";
RL Submitted (Feb-2005) to the EMBL/Genbank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
EMBL: AY921578; AAX84029.1; -; Genomic_DNA.
DB SRR: Q4ZE94; 1-561.
GO GO:0005576; C:extracellular region; IEA.
GO GO:0009405; P:pathogenesis; IEA.
DR InterPro: IPR03896; Antirax toxinB.
DR Pfam: PF03495; BimA_toxI.
DR PRINTS: PR01391; BINARYTOXINB.
KW Plasmid.
KW NON TER
SQ SEQUENCE 561 AA; 62846 MW; 88D32AACDF5955FA CRC64;

Query Match 72.8%; Score 2744; DB 2; Length 561;
Beet Local Similarity 96.3%; Pred. No. 1.4e-124;
Matches 543; Conservative 4; Mismatches 11; Indels 6; Gaps 3;

QY 175 VPDRLNDGIPSLSEVEGTYVVKKRTFLSPWISNIHEKGLTKYKSSPEKWSSTADSY 234
DB 1 VPDRLNDGIPSLSEVEGTYVVKKRTFLSPWISNIHEKGLTKYKSSPEKWSSTADSY 60
QY 225 DPEKVTGRIDKNSPEARHPVLAAYPIVHVMENIILSKNEQSTQNTDSQRTISKNTS 294
```



```

DB      61 DEKVTGRIDKNVSPARHPVLAAPYIVHDMENILLSKNEOSTONTOSQRTIKNTS 120
QY      295 TSDANTVGVSIAGYONGF---TGNITTSAGFSNNSSTVAIDHSLSLAGERTMAETWG 351
DB      121 TTR-THTSEVHGNAEVAHAFPDIGSV--SAGFSNNSSTVAIDHSLSLAGERTMAETWG 177
QY      352 LMTADTARLANINIRYVNTGAPRYNVLPTTSLVLGKQOTLATIKAVENOLSQLAPNNTY 411
DB      178 LMTADTARLANINIRYVNTGAPRYNVLPTTSLVLGKQOTLATIKAVENOLSQLAPNNTY 237
QY      412 BSKNLAPIALNADDDSSPTITNNVNOFLEKTKOLRLDTPDVGNIAITVYENGRVY 471
DB      238 PSKNLAPIALNADDDSSPTITNNVNOFLEKTKOLRLDTPDVGNIAITVYENGRVY 297
QY      472 DTGSNNSEVLPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKALKIAP 531
DB      298 DTGSNNSEVLPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKALKIAP 357
QY      532 GENEPNNGTQYQKDIITEPDFNFDQTSQNIKNQALAEIANTYIVLDKIKANNNILI 591
DB      358 GENEPNNGTQYQKDIITEPDFNFDQTSQNIKNQALAEIANTYIVLDKIKANNNILI 417
QY      592 RKRFRHYDNNNTAVGADSEVYKAEHREVINSSTEGLLNIDKDIRKILSGYIEIDTGS 651
DB      418 RKRFRHYDNNNTAVGADSEVYKAEHREVINSSTEGLLNIDKDIRKILSGYIEIDTGS 477
QY      652 LKEVINDRYDMLNLSLRQDKTFIDFKKYNDCLPLYISNPNKYVAVYATKENTIIINS 711
DB      478 LKEVINDRYDMLNLSLRQDKTFIDFKKYNDCLPLYISNPNKYVAVYATKENTIIINS 537
QY      712 ENGDTSTNGIKKILIFSKKGYEIG 735
DB      538 ENGDTSTNGIKKILIFSKKGYEIG 561

RESULT 5
Q46221_BACCE PRELIMINARY; PRT; 760 AA.
ID 046221_BACCE PRELIMINARY; PRT; 760 AA.
AC 046221;
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE Protective antigen.
GN ORFNames=BCE G9241_PBC218_0026;
OS Bacillus cereus G9241.
OC plasmid pBC218
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxId=269801;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15155910; DOI=10.1073/pnas.0402414101;
RA Hoffmayer A.R., Ravel J., Raiko D.A., Chapman G.D., Chute M.D.,
RA Maister C.K., De B.K., Sacchi C.T., Fitzgerald C., Mayer L.W.,
RA Maier M.C.J., Priest F.G., Barker M., Jiang L., Cer R.Z.,
RA Rilestone J., Peterson S.N., Weyant R.S., Galloway D.R., Read T.D.,
RA Popovic T., Fraser C.M.;
RT Identification of anthrax toxin genes in a Bacillus cereus associated
RT with an illness resembling inhalation anthrax."
RT Proc. Natl. Acad. Sci. U.S.A. 101:8449-8454(2004).
-1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivative License
CC -----
EMBL: AABK0100004; EAL15944.1; -; Genomic DNA.
DR GO: GO:0005576; C:extracellular region; IEA.
DR GO: GO:0009405; P:pathogenesis; IEA.
DR InterPro: IPR003896; Anthrax_toxinB.

```

```

DR      InterPro: IPR013145; PA14_rel.
DR      Pfam: PF03495; Binary_toxB, 1.
DR      Pfam: PF07691; PA14, 1.
DR      PRINTS: PR01391; BINARYTOXINB.
DR      KX plasmid.
SQ      SEQUENCE 760 AA; 85765 MW; 7488F5B5BF1F054 CRC64;

Query Match      56.8%; Score 2140; DB 2; Length 760;
Best Local Similarity 58.3%; Pred. No. 3,4e-95;
Matches 427; Conservative 118; Mismatches 177; Indels 10; Gaps 7;

QY      4 QENRLNSESSESGGLGYFSDLNFOAPMVVTSSTYGLSIPSSLEN-IPSENQYFOS 62
DB      33 QEDSLIQKETSQGLGYFSDLNFOAPMVVTSSTYGLSIPSSLEN-IPSENQYFOS 92
QY      63 AIWSGFIKYKKSDEYTFATSDNHYTMVVDQEVINKASNKIRLEKRLYOIKYOYR 122
DB      93 AIWTFGVNVEKQGEYFTSDNNHVTISVNDQEVNMRSTDHLKIMLEKQLYEITVYKOR 152
QY      123 ENPTEKGLDPKLYWDSQKKEVYSSDNQLPELKQSSNSRKRKSTAGPYVPRDNDG 182
DB      153 ESPSEKGIQFOLYWTTPDWTTEITPKSKLLPLPKAKSSKSTERSANNRREIVDEDNDG 212
QY      183 IPDSLVEGTVYDVKNNKRTFLSPWISINHEKKGLTYKSSPEKMSYASDPYDFEKYVGR 242
DB      213 IPDSLVEGTVYDVKNNKRTFLSPWISINHEKKGLTYKSSPEKMSYASDPYDFEKYVGR 272
QY      243 IDKNVSPARHPVLAAPYIVHDMENILLSKNEOSTONTOSQRTIKNTSRTD--A 299
DB      273 IDKQVLEARDPRIAPFVNTVDMEKIISKIEDKSKQGGSGQNTISGTSTSKHTYLE 332
QY      300 NTWGVSIAGYONGFNGITTSAGFSNNSSTVAIDHSLSLAGERTMAETGINTADTAR 359
DB      333 EHADVSPHSLPD-FGGSV--SAGISNNSSTVYIDTSSSTNLEKSSWSETLVGNTADAAK 389
QY      360 LMANIRYVNTGAPRYNVLPTTSLVLGKQOTLATIKAVENOLSQLAPNNTYYSKMLAPI 419
DB      390 LMANIRYVNTGAPRYNVLPTTSLVLGKQOTLATIKAVENOLSQLAPNNTYYSKMLAPI 449
QY      420 ALNAODPFSSPTITNNVNOFLEKTKOLRLDTPDVGNIAITVYENGRVYVDTGSNWS 479
DB      450 SINAODPFSSPTITNNVNOFLEKTKOLRLDTPDVGNIAITVYENGRVYVDTGSNWS 509
QY      480 VLPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKALKIAPFPNPN 539
DB      510 FLPOIETTARIILQSDNNDSTERRVAAVDPNDFLEMTKPDVTLKALKIAPFKENRQ 569
QY      540 LQYQKDIITEPDFNFDQTSQNIKNQALAEIANTYIVLDKIKANNNILIDKGFHYD 599
DB      570 LKYEKDIISFEFVFKETNQNIOOLKGRKFNITBALDNIOLANKMILVVDANFRY- 628
QY      600 RNNIAGADESVYKAEHREVINSSTEGLLNIDKDIRKILSGYIEIDTGLKEVINDR 659
DB      629 VDGKRYGADWDPIKTHQGVIKSESATLKIDKNTTMMAGLIEVERSGKGLVNSR 688
QY      660 YDMNLNLSLRQDKTFIDFKKYNDCLPLYISNPNKYVAVYATKENTIIINSNGTSTN 719
DB      689 YSDIDFSLIKGTVLDFKKNNDNTPRINN-GETVNIYAITKERNKLAIEGRPNIN 747
QY      720 GIKK-ILIFSK 730
DB      748 DSEKGLVLSLK 759

RESULT 6
Q46221_CLOPE PRELIMINARY; PRT; 875 AA.
ID 046221_CLOPE PRELIMINARY; PRT; 875 AA.
AC 046221;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE Iota toxin component Ib precursor.
OS Clostridium perfringens.

```

```
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OK NCBI_TaxID=1502;
RN
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=NCIB 10748;
RC MEDLINE=94041637; PubMed=8225592;
RA Perelle S., Gilbert M., Boquet P., Popoff M.R.;
RT "Characterization of Clostridium perfringens iota toxin genes and
RT expression in Escherichia coli."
RT Infect. Immun. 61:5147-5156(1993).
RN
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=NCIB 10748;
RC
RC Popoff M.R.;
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL, X73562; CA51960.1; -; Genomic_DNA.
DR PIR, I40862; I40862.
DR HSSP, P13423; IACC.
DR GO, GO:0005576; C:extracellular region; IEA.
DR GO, GO:0009405; P:pathogenesis; IEA.
DR InterPro, IPR003896; Anthrax_toxinB.
DR InterPro, IPR011658; PA14.
DR InterPro, IPR013145; PA14_rel.
DR Pfam, PF03495; PA14_rel.
DR Pfam, PF07691; PA14; 1.
DR PRINTS, PR01391; BINARYTOXINB.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 875 AA; 98469 MW; C9AE092CD818921 CRC64;

Query Match 26.1%; Score 984; DB 2; Length 875;
Best Local Similarity 34.2%; Pred. No. 3.3e-39;
Matches 275; Conservative 125; Mismatches 268; Indels 136; Gaps 32;

QY 1 EVKQENRLNESSESSOGILGYFSDLNFOAPMVYTSSTTGDLSPSELENIPSE-NQY 59
DB 35 DTNGEBEITNENTLNSNGMGVYPADBEHKDLEMAPIKNGDLKFEKKVVDLITEDNS 94
QY 60 FQSAWSEFIAKKKSDETFTATSDNHTVMVDDQEVINKASNSKIRLEKRLYQIKIQ 119
DB 95 IKSIMTGRILIPSBGGEYLSTDR-NDVLMQINAGDLAK--TLKVNKKKQAAVNIIRLE 150
QY 120 YQREN-----PTEKGLDFKLWTDGQNKKEVYISDNLQPLKOKSSNSRKKRST 170
DB 151 IQDKNLGSLDNLVSP-----KLW-ELANGKTVIPEENLFFRDYSKIDEND----- 195
QY 171 AGPTVP-----DRNDGIPDLSLEVEGYTVDVYKNTFLSPWISNIH 211
DB 196 --PPIPNNNFPDVRPFAAMEDEDLDTQNDNIPDAVEKKGYIT----KQSIAMVKNDS 249
QY 212 EKKGLTKYKSSPEKKSTASDPYSDPEKVTGRIDKXVSPAPRPVLAAPYIVHVMENITL 271
DB 250 E-QGYKAYVSSYLESTAGDPYTDYOKASGSIDKAIKLEAPDPLVAAVPAVVGEMENLI 308
QY 272 SKNEQOSTONTDSQRTISKNTSTSR-DANTVGVSISAGYONGFTNGITTSAGFSNSNS 330
DB 309 STNEIAS--DQGTIVSRATNTSKTDNITVGVSISAGYONGFTNGITTSISHTTNDST 364
QY 331 TVAIDHSISLAGERTWAEWTGLNTADTARLANNIRVYTGTAPIYVNLPTTSLVLGKQOT 390
DB 365 AVQDSN-----GE-SWNTGLSINKGESAVINANVRVYTGTAAPKVPPTNLVL--DDET 417
QY 391 LATIYAKENQLSQTIAAPNNYPSKILAPIALAAQDDFSTPTTMYNQGLELEKTKQRL 450
DB 418 LATIYAKQDQIQNNLSIPNETYPPKGLSPLALNTMQFNARLIPINYDQIKLKDQSKQIDL 477
QY 451 DTDDYGNATATYNFENGRVRYVDYTGNSWSEVLPIQIDETARIIIFLNGDLNLYERRIAAVNP 510
```

```
DB 478 ETTQVSSNGVYTKN-SQGGI-ITBEGNSMNSVYSQIDSVSASITILD-TSGQTFERRVAAEQ 534
QY 511 SDPLETTKPDNTLKBALKIAPGNEPENGLOI-QGKDITE--FDPNFPQOSTSONIKNOLA 567
DB 535 GNPEDKT-PEITIGSAIKKAFSAATK-NGELLVFNQIPIDESCVLEIFPDNTSEIIKEQLK 592
QY 568 ELNATNYTVYDVKIKLNKKNMILIRDKRF--HYDR-NNIAVGADESIVYKAEAREVYNSST 624
DB 593 YLDDKKIYVNV--KLERGMNLIKVPSTFNNPBNP--ASMSNIDTKNQDGLQSV 646
QY 625 EGL-----LNIIDKDIRKILSGY-----IVEIEDTEGLKEVINDRYDMANI 665
DB 647 NKLGETKILIPMKLKYKRYVFSGYSDKSTNSITVNIKSKQKTDVYLPKEDYK 706
QY 666 S-----SLRQDGKTFIDPKKYN--KLPLYISNPNYK-----N 697
DB 707 SYEFTTGKSDSIEITLTSVGFILDMLSITEINSPLEIKBEPIKVPDQEIILDAHK 766
QY 698 VYAVTKENTLIINPSENGDSTNGI 721
DB 767 YYADIKLDT-----NTGNTYIDGI 785

RESULT 7
ID O32739_CLODI PRELIMINARY; PRT; 876 AA.
AC O32739;
DT 01-JAN-1998, integrated into UniProtKB/TrEMBL.
DT 01-JAN-1998, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE ADP-ribosyltransferase.
GN Name=cotB;
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OK NCBI_TaxID=1496;
RN
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=CD196;
RC MEDLINE=97230316; PubMed=9119480;
RA Perelle S., Gilbert M., Bourlioux P., Corthier G., Popoff M.R.;
RT "Production of a complete binary toxin (actin-specific ADP-
RT ribosyltransferase) by Clostridium difficile CD196."
RL Infect. Immun. 65:1402-1407(1997).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL, I76081; AAB67305.1; -; Genomic_DNA.
DR HSSP, P13423; IACC.
DR GO, GO:0005576; C:extracellular region; IEA.
DR GO, GO:0016740; P:transferase activity; IEA.
DR GO, GO:0009405; P:pathogenesis; IEA.
DR InterPro, IPR003896; Anthrax_toxinB.
DR InterPro, IPR011658; PA14.
DR InterPro, IPR013145; PA14_rel.
DR Pfam, PF03495; Binary_toxB; 1.
DR Pfam, PF07691; PA14; 1.
DR PRINTS, PR01391; BINARYTOXINB.
KW Transferase.
SQ SEQUENCE 876 AA; 98797 MW; 25E06E2D45CE2B3B CRC64;

Query Match 25.6%; Score 966.5; DB 2; Length 876;
Best Local Similarity 33.8%; Pred. No. 2.3e-38;
Matches 266; Conservative 137; Mismatches 274; Indels 111; Gaps 31;

QY 4 QENRLNESSESSOGILGYFSDLNFOAPMVYTSSTTGDLSPSELEN-IPSENYQFOS 62
DB 40 KKEKIVNEDILIPNNGLMKYFSDPEHKDKLMAPIKQDNLKFBEKKVVDLKDQSDVKS 99
QY 63 AIWSEFIAKKKSDETFTATSDNHTVMVDDQEVINKASNSKIRLEKRLYQIKIQYOR 122
```

```
Db 100 IRWTGRIIPSKDGEYTLSTDRD-VLMQVNTBSTI---SNTLKNVKKKGEYKRIELOD 155
Qy 123 EN--PTEKGLDFLKYTDSQNKKEVISDNLOPELKQSSNSRKRSTAGTVP---- 176
Db 156 KNLGSIIDNLSSPPLYM-ELDKMKKIIPEENLFLRDY----SNIEKD-----DPIFPNNF 205
Qy 177 -----DDNDGIPDLSLEVEGYTYDVKNKRTFLSMISNHEKKGLTYKKS 222
Db 206 FDPKMSDWEDEDLTDNDNIPDSYERNNGTYI---KDLIAVWEDSPAE-QGYKXYVSN 260
Qy 223 PEKMTASDPYSPDEFKVTGRIDKNVSPEARHPLVAAPYIVHVMENIISKNDSOSTONT 282
Db 261 YLESNTASGPDTYTYEKAASSGFDKAITKEARDPLVAAPYIVGMEKLIISTNEHAST--- 317
Qy 283 DSQRTIRISKNSTSR-DANTVGVISAGYONGFTGNITTSAGSNNSTVAIDHSLSLA 341
Db 318 -DQKTVSRAITNSKTESNTAGVSVVNGONGFTAVNTVSHTTDNTSTAVQSN----- 371
Qy 342 GERTMAETGLNADTARLANIRYNTGTAPLYNVLPTTSLVLGKNQTLATIKAKENOL 401
Db 372 GE-SMWTGSLINGESAYINANVRYNTGTAPMYKVTPTNLVL-DGDTLSTIKAEENOI 429
Qy 402 SQTIAENNYPSKNLAPIALNAODDFSTPTITANNQPLELEKTKQLRLDTPQVGNIAI 461
Db 430 GNNLSFGDTYPRKGLSPLALNTMDQSSRLIPINYDQLKLDAGKQIKLETTOVSGNFGT 489
Qy 462 YNEENGRVAVDTGSSNMEVLPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPM 521
Db 490 KN-SSGOI-VTEGNSMSDYISQIDSISASIIIDTEN-ESYERRVTAKNLQDPEDKT-PEL 545
Qy 522 TLKALKIAGFENEPNGNLOYGKDIITE--FDENFDQOTSQNIKNQLABINATNIYTVLD 579
Db 546 TIGEAIEKAFATKQGLLYFNDIPIDESCEVELIFDNTANKIKDSLKTISDKKIYV-- 603
Qy 580 KIKLNAKNLILDRKRFHY---DRNNIAGADESVYKAEHRYVNSTEGL----- 627
Db 604 --KLERGMNLIKPTPTFTNPDYNNYP--STWSNVNTTNQDGLQSSANKLNGETKIKP 659
Qy 628 LNLNDKDIRKILSGY-----IVEI-----EDTEGL-----KEVINDRYDML 663
Db 660 MSBLKPKRYKRVFSGYKDPJLTSNSIIYKIKAKEKNDYLVPEGOYTKFSYEPFETEKDS 719
Qy 664 NIS-SLRQDKTIFIDPKYND--KLPLYISNPNYK-----VNTVAVTKENTINPSEN 713
Db 720 NIEITLIGSGTYLDMLSITEINLSTPEIIDBEVKIPTDQEIINDAHKIYFADLNFPNS-T 778
Qy 714 GDSSTNGI 721
Db 779 GNTYINGM 786

RESULT 8
09KH41_CLODI PRELIMINARY; PRT; 876 AA.
AC 09KH41;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Cdb.
GN Name=cdb;
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_Taxid=1496;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CCUG 20309;
RX MEDLINE=21604263; PubMed=11761709;
RA Chang S.Y., Song K.P.;
RT "ADP-ribosylating binary toxin genes of Clostridium difficile strain
CCUG 20309."
RL DNA Seq. 12,115-120(2001).
```

```
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivative License
CC -----
CC EMBL; AF271719; AAF81761.1; -, Genomic_DNA.
CC HSSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax_toxinB.
DR InterPro; IPR011658; PA14.
DR InterPro; IPR013145; PA14_rel.
DR Pfam; PF03495; Binary_toxib.
DR Pfam; PF07691; PA14; I.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 876 AA; 98793 MW; 366D62F352E745A5 CRC64;

Query Match 25.6%; Score 964.5; DB 2; Length 876;
Best Local Similarity 33.6%; Pred. No. 2,9e-38;
Matches 265; Conservative 138; Mismatches 274; Indels 111; Gaps 31;

Qy 4 QENRLANESSESSQGLGYFSPDLNFOAMVNTSSTTGDLISPSLEEN-IPSENGYFOS 62
Db 40 KKKEIVNEDILPNNGLMGYFTDEHPKDLKMAPIKDNLKEBKVDKLDKQSDVKS 99
Qy 63 AMSGFIRKYKSDYETFPATSDNHYTMVYDDQEVINKASNKRIRLEKGLYQIKIYOYR 122
Db 100 IRWTGRIIPSKDGEYTLSTDRD-VLMQVNTBSTI---SNTLKNVKKKGEYKRIELOD 155
Qy 123 EN--PTEKGLDFLKYTDSQNKKEVISDNLOPELKQSSNSRKRSTAGTVP---- 176
Db 156 KNLGSIIDNLSSPPLYM-ELDKMKKIIPEENLFLRDY----SNIEKD-----DPIFPNNF 205
Qy 177 -----DDNDGIPDLSLEVEGYTYDVKNKRTFLSMISNHEKKGLTYKKS 222
Db 206 FDPKMSDWEDEDLTDNDNIPDSYERNNGTYI---KDLIAVWEDSPAE-QGYKXYVSN 260
Qy 223 PEKMTASDPYSPDEFKVTGRIDKNVSPEARHPLVAAPYIVHVMENIISKNDSOSTONT 282
Db 261 YLESNTASGPDTYTYEKAASSGFDKAITKEARDPLVAAPYIVGMEKLIISTNEHAST--- 317
Qy 283 DSQRTIRISKNSTSR-DANTVGVISAGYONGFTGNITTSAGSNNSTVAIDHSLSLA 341
Db 318 -DQKTVSRAITNSKTESNTAGVSVVNGONGFTAVNTVSHTTDNTSTAVQSN----- 371
Qy 342 GERTMAETGLNADTARLANIRYNTGTAPLYNVLPTTSLVLGKNQTLATIKAKENOL 401
Db 372 GE-SMWTGSLINGESAYINANVRYNTGTAPMYKVTPTNLVL-DGDTLSTIKAEENOI 429
Qy 402 SQTIAENNYPSKNLAPIALNAODDFSTPTITANNQPLELEKTKQLRLDTPQVGNIAI 461
Db 430 GNNLSFGDTYPRKGLSPLALNTMDQSSRLIPINYDQLKLDAGKQIKLETTOVSGNFGT 489
Qy 462 YNEENGRVAVDTGSSNMEVLPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPM 521
Db 490 KN-SSGOI-VTEGNSMSDYISQIDSISASIIIDTEN-ESYERRVTAKNLQDPEDKT-PEL 545
Qy 522 TLKALKIAGFENEPNGNLOYGKDIITE--FDENFDQOTSQNIKNQLABINATNIYTVLD 579
Db 546 TIGEAIEKAFATKQGLLYFNDIPIDESCEVELIFDNTANKIKDSLKTISDKKIYV-- 603
Qy 580 KIKLNAKNLILDRKRFHY---DRNNIAGADESVYKAEHRYVNSTEGL----- 627
Db 604 --KLERGMNLIKPTPTFTNPDYNNYP--STWSNVNTTNQDGLQSSANKLNGETKIKP 659
Qy 628 LNLNDKDIRKILSGY-----IVEI-----EDTEGL-----KEVINDRYDML 663
Db 660 MSBLKPKRYKRVFSGYKDPJLTSNSIIYKIKAKEKNDYLVPEGOYTKFSYEPFETEKDS 719
Qy 664 NIS-SLRQDKTIFIDPKYND--KLPLYISNPNYK-----VNTVAVTKENTINPSEN 713
Db 720 NIEITLIGSGTYLDMLSITEINLSTPEIIDBEVKIPTDQEIINDAHKIYFADLNFPNS-T 778
Qy 714 GDSSTNGI 721
```

Db 779 GNTYINGM 786

RESULT 9

ID 006498_9MOLU PRELIMINARY; PRT; 879 AA.

AC 006498;

DT 01-JUL-1997, integrated into UniProtKB/TrEMBL.

DT 01-JUL-1997, sequence version 1.

DT 07-FEB-2006, entry version 23.

DE Sb component.

GN Name=sbs;

OS Clostridium spiroforme.

OC Bacteria; Firmicutes; Mollicutes.

OX NCBI_TaxID=29348;

RN [1]

RC NUCLEOTIDE SEQUENCE.

RA Gilbert M., Perelle S., Daube G., Popoff M.R.;

RT "Clostridium spiroforme toxin genes are related to C. perfringens iota toxin genes but have a different genomic localization.";

RT Syst. Appl. Microbiol. 20:337-347(1997).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NonDerivs License

EMBL: X97969; CA66612.1; -; Genomic_DNA.

DR HSP; P13423; IACC.

DR GO; GO:0005576; C:extracellular region; IEA.

DR GO; GO:0009405; P:pathogenesis; IEA.

DR InterPro; IPR003896; Anthrax_toxinb.

DR InterPro; IPR013145; PA14.

DR Pfam; PF03495; BimA4_rel.

DR Pfam; PF07691; PA14; 1.

DR PRINTS; PR01391; BINARYTOXINB.

DR SEQUENCE 879 AA; 98739 MW; 4065ACB8E05BA01 CRC64;

Query Match 25.5%; Score 960.5; DB 2; Length 879;

Best Local Similarity 30.2%; Pred. No. 4,5e-38;

Matches 272; Conservative 146; Mismatches 248; Indels 235; Gaps 34;

QY 4 QENRLNSESSESGGLGYPFDLNPQAMVVTSSFTGDLSPSSELENIPENO-YFQS 62

Db 36 QEVETNTEKTVSSNGIMGYFADHFQDELAPYKNGELKEKKNVEKLLTEKNTKS 95

QY 63 AIWSGFIKVKSDSEYFATSDNHYTMVVDQEVINKASNKIRLEKRLYQIKIYQYR 122

Db 96 IRMTGRITIPSKQGEYTLSTDKN-VLMQINAE--GEIANTLKVMIMIKQGEYSIRIEQ- 150

QY 123 ENPTREKGLDF-----KLYWDSQNKKEVISDNILQPELKQKSN----- 162

Db 151 ---PKDIGYVDLSPPKLYW-ELMGDKTLIPKMLFLADYSKIDENDPFIPKDNFFDLK 205

QY 163 ---SRKKRSTSA-GPTVPDRDNDGIPDSLEVEGYTVDKKRFPLSPWISNHEKKGLTGY 219

Db 206 LKSRARLARGWGDDELDLDNNTIPAYEKNGTIT---KDSIAYKW-EDSPAQGGYKKT 260

QY 220 KSSPEKWSYASDPYSDPEFYKTRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDST 279

Db 261 LSSYLESNATGAPYDYQKASGSPKAIQAEARDPLVAAYPVGVGMELIISTHEHAST 320

QY 280 QNTDSQRTISKNSTSR-DANTVGVISAGYQNGFTGNTTSAGFSNSNSNTVAIDHSL 338

Db 321 ---QGGKVSHTNTSKTDANTAGVAINIAYQNGFTGSIITN--YSHTEENSTAVQNS- 373

QY 339 SIAGRTAETMGMLADTARLANRINYNTGAPTYVNLPTTSIVLGNQGLATIKAE 398

Db 374 --NGR-SWNTSLSINKSBAVINANRYNTGAPYKATPTTNLVL-DGDTLVYTIKAD 429

QY 399 NQLSQILAPNNYPSKNIAPIALNADDDPSSTPTITWYVQPLEKTKQRLADTDQVYGN 458

Db 430 NQIGNNLSNENYPPKKGSLPLALNTMDGSSRLIPINTYQKKLDAKGQIKLETTQVSGN 489

QY 459 IATYFENGVRVDT-GSMSEVLPOIQTETARIIFN-GKDLNVERRIAANPSDPLET 516

Db 490 ---YGIKNSQGIITTEGNSWSDYISQIDSLASIIIDGSD--VPERVYAKDSNSPBDK 544

QY 517 TKPDMTLKALKIAFGFENPGNIQYQCKDITE--PDFNPDQOSTONIKQALAEINATNI 574

Db 545 T-PVLTIGEALEKAFGATKNGEILYFNQMPIDESCVELIPGANTANLIKERLINALNDKKI 603

QY 575 YTVLDKTIQANKNMILRPD----- 593

Db 604 YNV---QLERGKMLIKXTSTYFNNPDGYNPPSSWSNVDSNODGLQNAANKLSGETYKI 659

QY 594 -----KRFHY----- 598

Db 660 VIPMSKLNPKRYVPSGLKNSSTSNPTTNVIAKQKTYNLVSENDYKKTSTYERETIGR 719

QY 599 DRNN-----IAGDAESVVKAREV-----IN 621

Db 720 DASNIEITLTSSGTLFLDNLSTELNSTPEILKEPDIKVPEDQELI-DAHKKYADLSFN 778

QY 622 SSTEGLLIN-----IDKDIRKILSGIYVEIE-DTEGLKEVINDRYDMANISLRQDG 672

Db 779 QSTANYVLDGLYFEPYQTNKEVLDYIQKYKVEATLEYSGFQDIGNKDELKNYTGDSNQ 838

QY 673 KT-FIDFKYNDKLPLYISNPNY---KVNYYAVTKENTIIINPSENGDSTNGIKILIF 727

Db 839 KINYVNFSTY-----FTSGENWMPYKGLRIYALTPEN-----KELLVL 876

QY 728 S 728

Db 877 S 877

RESULT 10

ID 086171_CLOBO PRELIMINARY; PRT; 721 AA.

AC 086171-

DT 01-NOV-1998, integrated into UniProtKB/TrEMBL.

DT 01-NOV-1998, sequence version 1.

DT 07-FEB-2006, entry version 19.

DE C2 toxin (Component-II).

OS Clostridium botulinum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.

OX NCBI_TaxID=1491;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN-type C;

RX MEDLINE=96184657; PubMed=8645309; DOI=10.1006/birc.1996.0409;

RA Fujii N., Kubota T., Shirakawa S., Kimura K., Ohishi I., Morishih K., Isogai E., Isogai H.;

RT "Characterization of component-I gene of botulinum C2 toxin and PCR detection of its gene in clostridial species.";

RT Biochem. Biophys. Res. Commun. 220:353-359(1996).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN-type C;

RX MEDLINE=96323874; PubMed=9659689; DOI=10.1016/S0378-1135(98)00195-3;

RA Kimura K., Kubota T., Ohishi I., Isogai H., Isogai E., Fujii N.;

RT "The gene for component-II of botulinum C2 toxin.";

RT Vet. Microbiol. 62:27-34(1998).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NonDerivs License

EMBL: D88982; BAA32537.1; -; Genomic_DNA.

DR HSP; P13423; IACC.

DR GO; GO:0005576; C:extracellular region; IEA.

DR GO; GO:0009405; P:pathogenesis; IEA.

DR InterPro; IPR003896; Anthrax_toxinb.

DR InterPro; IPR011658; PA14.

DR InterPro; IPR013145; PA14_rel.

```
DR  Pfam; PF03495; Binary_toxB; 1.
DR  Pfam; PF07691; PA14; 1.
DR  PRINTS; PR01391; BINARYTOXB.
SQ  SEQUENCE 721 AA; 80516 MW;
```

Query Match	22.0%	Score 829;	DB 2;	Length 721;
Best Local Similarity	35.4%;	Pred. No. 8e-35/		
Matches 224; Conservative	120;	Mismatches 224;	Indels 64;	Gaps 26

Qy	3	KOEHRRLNSES--SSOGLGCTYFSDNLFOAPMWTASTGDLISPSSELENIPEBNQYF	60
Db	5	KFENSVAKNKMYTINGLMGYTEN-DFFNLNTISPLDGLNLFKSKDINSILG-NKII	62
Qy	61	QSAIWSGFIVKKSDEYTFATSAUN-HVTMWVDDQEVINKASN-SNKIRLEKRLYQIKI	118
Db	63	KSARWILKIPRSITGEYILSTNSPNCREVL--NGEINLSTLNTSNVNLIGQNYVDIRI	119
Qy	119	-QYORENPTKEGLD-FLKYWTDSQNKKEVISDSNOLPELKQKSNKSKKSTAGPIV-	175
Db	120	EQLMSENOQLKNYGIKLYWETSJIIKEIPESEVL---LKPNSNTNEKKRPINNTLF	175
Qy	176	-----PDRNDGIPDSLEVEGYVDVQNKRTFLSPWISNHEKKGLTKYKSP	225
Db	176	SNALKANANARDTRDGLPEDEMEINGTYVMNQKAVAMODKRAAN-----GYKKVSNFK	230
Qy	226	WSTASDYPDSFEKTYGRIDKNVSPARHPVAAPYVHVDMENIILSKNEDOSTONTDSQ	285
Db	231	PCTANDEYTFEKFSGGQIDPEVSWARPMISAPYIVGOMERLWVSKSE---TITDSD	286
Qy	286	TRTISKSTS-TSRBANTGVGISAGYQNGFTGNI-----TTSAGFSNSSTVAIDHSL	338
Db	287	TKMSKSTSHSSTININIVGAESGSLQ--LAGGIFPVPSMASANASTHWQTSIVDQT-	343
Qy	339	SLAGERWAEWTMGINTADTARLANINIRVYNTGTAPIVNLTPTSVLGKQNTATIRAKE	398
Db	344	--TGE-SFSQGLSINTGESAYINPIRIRVYNTGTAPVNVVPTTTIVIDK-QSVAITIKGE	399
Qy	399	NOLSOIILAPNNYYSKULAPIALNAQDDFSSPTPMYANTQGLEKTYQQLRLDTPQVYGN	458
Db	400	SLIGDYINPGGTYPYIGEPMPALWMTQPFSSRLPIPNINOLKSIDNGETVWLSTSQFTGN	459
Qy	459	IATYFENGRGRVVTGSNWSSEVLPOIOETTRARI--FNGKOLNVERILIAVNPSPDET	516
Db	460	PAKYN-SNGNIVYD-GNNMGPYLGTIKSTTASLTLSFSGQTTQYA--VVAPNFSDPEDK	514
Qy	517	TKEDMTLKEALKLAFGNBPNGNLQYQOKDIT--EPDFNDQOOTSONIKNQULAEINATN	573
Db	515	T-PLTLEQLLVKPALEKKNQKGFYFPHGLEISKNEKIQVFLDSNTNND FENQLKNTADXD	573
Qy	574	IYYVLDIKILAKANKNILIRDKRFHYDRNNIAV	605
Db	574	IMHCI--IKGN--MNLVKTITPEKNTSSINI	601

RP NUCLEOTIDE SEQUENCE.
RA Van Rie J.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AJ872073; CA13278.1; -; Genomic DNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax_toxinb.
DR InterPro; IPR011658; PA14.
DR InterPro; IPR013145; PA14_rel.
DR Pfam; PF03495; Binary_toxb_1.
DR Pfam; PF07691; PA14_1.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 785 AA; 87619 MW; F8A87A37D91A4D1E CRC64;

Query Match	20.8%	Score 786	DB 2	Length 785
Best Local Similarity	30.7%	Pred. No. 1	1e-29	
Matches 236	Conservative 119	Mismatches 239	Indels 114	Gaps 27

```

0Y 4 QENLILNSESSESSGGLLGGYFSDILNCPAPWVWISSTGDISBSSELE--NIPSNQYFOS 62
Db 42 QENQ---EKVDRKRGGLGYFFKGFNFHNLFLPAPTRNTLIIYQQYTDPLDKKQSGYOS 98
0Y 63 AIMSGLIKVKSDEYFATSDADNHVTMWVDDQEVINFASSNKLRLKGRLYOIKTOYOR 122
Db 99 IRMGLIGSRETGDPFNLLDDKNAILEIDOKTISHGQKQVYHLEKGLVPIKIEYOS 158
0Y 123 ENFPEKG---LDBKLYWTDSONKGEVYISSDNQLPELKQKSSNRKGSSTAG----- 172
Db 159 DOIIVNRORSKIFKEKPKLFKVDSONQOSHOUDELNPFNFNKEETOQFLEKASKTNIPLOQM 218
0Y 173 ---PTVEDRNDGJPDLSJEVGYVD---VKNRKPTLSPWISNIEHKGLTKYKSSPEK 225
Db 219 KRDEADATDDBDSJPDIMEENGTTIOUKVAVKMDSDFA-----KQYTFVSNPFD 269
0Y 226 WSTASDEYSDPEKVTGRIDKNVSEBAPPLVAAPYIVHDMENIISLKNEQSTONTDSQ 285
Db 270 SHYVGDEPYDYEKKAARLDLANAKETNPPLVAAPSVNVLKYLISPNEDLS----- 322
0Y 286 TRTISKNTSTSRD--ANTVGVYSISAGYONGTGNITTSAGSNS--NSSTVALDHSLSLAG 342
Db 323 -NSVESHSSTWSTYNTGEGASIEAG-----SGPLGISYGVSAVYQHSSETVA----- 367
0Y 343 ERTAEAFMG-----LNTADTARILANRILRYVNTGTAPRYNVLPTTSYLVYKGNQTLATIKAXE 398
Db 368 -KENGSTGTGTSOPNTIASAGYLANNVKXNNVGALIEVAPPTGTFVL--DNDYATITAKS 425
0Y 399 NQLSQILAPNNYYSKMLAPIALNAODFESTPTITANNYNOFLELEKTKOULRDTQOYVN 458
Db 426 NSTALSTSPGSSYKKGONGAIINTMDDFSHPTLTKQOLDQFNNKPLMLEFNQADG-- 484
0Y 459 IATYFENGRVYDTGSNMSEVLFOIOETRARIIFNKQDLNLYERRAAYANPSPDELETK 518
Db 485 --VYKIDTSGNIVTGEKNGVITQIOIAKTASIIIVDGE--GVSEKRYVAADQYNDPEKXT- 540
0Y 519 PDMFLKALKIAF--GNPENGNLQYOGKDIETEDF--NPDQOOSQIKNQLAELNA--T 572
Db 541 PSLLEKALKIAPPEEIKERDGLLYNDKPIYESSVNTYLDENTAKEVKQOLNDITGKFK 600
0Y 573 NIYVVLKIKLANAMNILLIRDRKFHYD-----RNNIVAGADESVKAEHVEVINSSTEGEL 627
Db 601 DVKQLPD-VLUTPRMNTTILKATIL--YQGAADGSSPTDVIGISSPLGEWAFFPDINNVEGN 658
0Y 628 LNLNDIKRLKSLGYIVEIDTBGLKEVINDRYDMNLISLRODGKTFIDFKKYNKPL 687
Db 659 TGKQYOLOSKKQGY-----YGMALAS-----PVSNSML-- 688
0Y 688 YISPNPKXAVVAYTKENTLIINPEN--GPISTNGI--KILIFSKGYE 733
Db 689 ---KRNQYIYISMSIKADAGEPPTVYVMULNGLIYDKKLLSNGYQ 733

```

```
RESULT 12
Q5JZ26_9BACL PRELIMINARY; PRT; 871 AA.
ID Q5JZ26_9BACL
AC Q5JZ26
DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.
DT 15-FEB-2005, sequence version 1.
DT 21-FEB-2006, entry version 6.
DE IspA protein.
OS Brevibacillus laterosporus.
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.
CC NCBI_TaxId=1465;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=16229973; DOI=10.1016/j.gene.2005.07.033;
RA Mestrovic N., Randig O., Abad P., Plohl M., Castagnone-Sereno P.;
RT "Conserved and variable domains in satellite DNAs of mitotic
pathogenetic root-knot nematode species.";
RL Gene 362:44-50(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Van Rie J.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
-----
DR EMBL; AJ871923; CAI40767.1; -; Genomic DNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax_toxinB.
DR InterPro; IPR01658; PA14.
DR InterPro; IPR013145; PA14_rel.
DR Pfam; PF03495; B:binary_toxB.
DR Pfam; PF07691; PA14_1.
DR PRINTS; PR01391; BINARYTOXINB.
DR SEQUENCE 871 AA; 97391 MW; 82P5EAA16763FDC5 CRC64;

Query Match 20.6%; Score 776; DB 2; Length 871;
Best Local Similarity 31.5%; Pred. No. 3.7e-29;
Matches 237; Conservative 123; Mismatches 296; Indels 96; Gaps 32;

QY 11 ESESSQGLGYFSDNLFQAPMVVTSSTTGDLSPSSLEN--IPSENQYQSAIMNGF 68
DB 47 DNQIDREGSLGYFFGKDFN-DLTFAPRDTLLYYDQTAATLVQKQEHYSIRWIGL 105
QY 69 IKVKSDEYTFATSDNHNVTMMVDDQEVINKASNSNKIRLEKGRLYQIKIYQYR----- 122
DB 106 IQSSATGDFTFGLSDENALIELDKVISEKNNQSVHLEKGLVQIKIEYQSDALHI 165
QY 123 ENPTEKGLDFKLYWTDSONKKEVISDNILQLELKOKSSNRKSTAG----- 172
DB 166 DNKIFK--ELKIFKIDSQNHSGOVQODELRNPEFNKKEQVLEKFKASKNMLFTQTKEDI 223
QY 173 PTVPRNDGTFDSELEVBGYTVDVANKRTFLSPWISNHEKGLKRYKSSPEKMTADP 232
DB 224 DEDTDTDSDSPDWVEENGTYT--IQNK--VAVKWQDSL-ASKGYQKFTSNPLEAHTVGP 278
QY 223 YSDPEKTYGRIDKNVSPEARHPLVAAYPIVHVMENIILSKEDQSTQNTDQTRTISKI 292
DB 279 YSDYRKARDMELSNAKETPNLVAAPSVNLSLEKVLISKEDSHSVESQ-----STN 334
QY 293 TSTSDANTVGVISAGYQNGFTGNITTSAGFSNSNSTVALIDHSLISLAGERTMAETWG- 351
DB 335 WSYT---NTEGVNVAAGW-----SGLGPFQVSVNQHSETVANWGSATNDGT 380
QY 352 -LNTADTARLNANIVYVNTGTAIPYVNLPTTSVLGKQNTLATITAKENQSLIAPNNY 410
DB 361 HINGBSAIVLNANVYNNVNGTAIYETKRTYFII-DGTTIGITAKETWTATLTLLPDGS 439
QY 411 YPSKMLAPIALNAODFSSTPYTMVNGP-LLEKTKQLRLDTPDQVYGIAATYENGAV 469
```

```
DB 440 YPEKKGKGIATINWDDFNSRPIPLNKEQLNTYLSNKKPILLETDOVSGKVAIKD-TNGNI 498
QY 470 RVDTSQMSSEVLPIQIETTAIIF-NGKDLNLYERRIAAVNPSDLETTKDMTLKEALK 528
DB 499 TI--AGDWNGITDEISAKTASIIYDNGNQMS--EKRYAAKQYTNBEDKT-PNLSYKEALK 553
QY 529 IAF--GFNEPFGNIOYQKDIITEPDPN--FPQOTSNIKNQALAINAT-----NIYVLAD 579
DB 554 LAYDEIEKKGGLFYNDQPIFEASVQSYVDYTRAKQIRKQNLNSTGSEFKQVKNLYDV-- 611
QY 580 KIKLNANKNIIIRDKRFHYD--RNNTAVGADSESVKAEHREYVNSTEGILLINDQIR 636
DB 612 --KLEPKNFTIKTSTL-YDGESEDNTKIG-----NMYYTYVYVNGGNTG-----KKQYR 657
QY 637 KILSGYVEI--EETEGIKXYINDRYDMLNITSLRQDKGTIFDPKXVDKLPYISNPN- 693
DB 658 SANRGAFTLESTESKKNKLNKI-DYVSL--YMKADSKSVVDIEIDGKQSIYTDNITL 713
QY 694 -----YKVNVAAYTKENTINP-SENGDSTYN 719
DB 714 DHVGQRNIIIVPLNENIINTISIKDGGQTN 745

RESULT 13
Q844J8_BACTU PRELIMINARY; PRT; 775 AA.
ID Q844J8_BACTU
AC Q844J8
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE VlpA.
GN Name=VlpA;
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxId=1428;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15357725; DOI=10.1111/j.1365-2672.2004.02365.x;
RA Shi Y., Xu W., Yuan M., Tang M., Chen J., Pang Y.;
RT "Expression of vlp1/vlp2 genes in Bacillus thuringiensis and the analysis of their signal peptides.";
RL J. Appl. Microbiol. 97:757-765(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
-----
DR EMBL; AY245547; AAO86514.1; -; Genomic DNA.
DR HSSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax_toxinB.
DR InterPro; IPR01658; PA14.
DR InterPro; IPR013145; PA14_rel.
DR Pfam; PF03495; B:binary_toxB.
DR Pfam; PF07691; PA14_1.
DR PRINTS; PR01391; BINARYTOXINB.
DR SEQUENCE 775 AA; 87111 MW; 416B25394361B731 CRC64;

Query Match 19.8%; Score 746; DB 2; Length 775;
Best Local Similarity 30.2%; Pred. No. 9e-28;
Matches 232; Conservative 128; Mismatches 292; Indels 116; Gaps 32;

QY 4 QENRLNESSSQGLGYFSDNLFQAPMVVTSSTTGDLSPSSLEN--IPSENQYQ 61
DB 42 QKNQ---QKENDRKGLGTYFFGKDF-SNLTMPFAPTRDNTLLIYDQTNKLLDKKQDQYQ 97
QY 62 SAISGFLIKVKSDEYTFATSDNHNVTMMVDDQEVINKASNSNKIRLEKGRLYQIKIYQ 121
DB 98 SIRWIGLQSKRGKDFYTNLSDEQALIEIDKILSNKGRKQVHLEKGLVPIKIRYQ 157
QY 122 RNPTEKGLD-----FRLYWTDSONKKEVISDNILQLELKQSSN----- 162
```

Db 158 SD--TKFNIDSKTPEKFKLTKIDSDQSQVYKRDLENPENKESREFLAKASKTNFEM 215
QY 163 SRKKSTSGAPTPPDNDGIPDLSLEVEGTTVDVKNKRTFLSPWISNIHEKGLTKYKS 222
Db 216 QKKRDRIDED---TDYDGDSDIPDLMEENGYT--IQNK--VAVWK--DDKFAQQCYVYKLSS 267
QY 223 PEKMSAPSDPYSPDFEYKVTGRIDKNVSPKAPVLAAYPIVHYVMENIILSKNDQSTQNT 282
Db 268 PYOAHVGDPTDWTWKAAGDIPKSNAAATNPVLAAPSPINVMKRTIKSKD-----SNL 322
QY 283 DSQRTISKNTSTSRDANTVGVISAGY--ONGFTGNITTSAGPSNSNSTVAIDHSLSLA 341
Db 323 SNSAEHNSNSYTY--ANSBGASIEAGFGPKGSPGV--SANT--QHTETVGS----- 370
QY 342 GERTW---AETWGLNTADTARLANIRYVNTGTAPVNTLPFTSLVGLKNOGLATITKAK 397
Db 371 ---WQNSKSNTEQFNSASAGYANAVHYNVGTGGVDAQPTTSPFL--QDSTIATITAK 425
QY 398 ENQLSQIILAPNNYPSKNIAPILANQDDSSPTPTMNNQPLEKTKQLRLDTQVYG 457
Db 426 SNATLSTISGDRYPASK--EGISLTKTMDDPNSHPTLNKPOLDAVINNEVYKINTQDTG 484
QY 458 NIATYFENGRAVVDGSMWSEVLPOIETTARIIENGKDLNVERRIAANPSDELETT 517
Db 485 RYGIIGVDG---KAEIGDRKSPIIDDEIKGRTAIIIDPADGRALERTIRAKOVKNEDKT 541
QY 518 KEDMTLKEALIKIAFG--FNEPENGNIQYQGD---ITBEDF-----FDQOTSQNIK 563
Db 542 -PSLTIKEGKTIAPESISEDKQGIIFYEYKNDGKVTYKQJLSEENIMPLDSDTSKEFE 600
QY 564 NQLAELANITVYLDKIKLANAKNIIILDKRPHYDRNNINAVGADSVYKAREVINS 623
Db 601 ROLSDGSAGKLY---DIKLTPTPMTTIR-----LATVTLFDDQFSAYPWNATWSP 649
QY 624 TEGSLT-----LNIIDKDIRK-----ILSGYIV--IEDTEGLKEVI---NDRYD 661
Db 650 KFGNRLGSLAIQESKTIIPKDKVKNPDYLTGIIKHDFTTDSLSGLVATTKDNFE 709
QY 662 MNISLSLRODGKTFIDFKKYNKDLPLYISNPNKYVNVYATKENTIN 709
Db 710 MNMGTISFSONSGSEFKFTIKTQ--NISG--DYILDSIQMKKNNDVN 755

RESULT 14
Y6163_BACAN STANDARD; PRT; 225 AA.
ID Y6163_BACAN STANDARD; PRT; 225 AA.
AD P13422; O8KXK2; O9X377;
DT 01-JAN-1990, integrated into UniProtKB/Swiss-Prot.
DT 05-JUL-2004, sequence version 3.
DT 07-MAR-2006, entry version 39.
DE Hypothetical protein pXOI-111/BXA0163/GBAA_pXOI_0163.
GN Ordered locus names=pXOI-111, BXA0163, GBAA_pXOI_0163;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1392;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=89172073; PubMed=3148491; DOI=10.1016/0378-1119(88)90439-8;
RA Welkos S.L., Lowe J.R., Eden-McCutchan F., Vockin M., Leppla S.H.,
Schmidt J.J.,
RT "Sequence and analysis of the DNA encoding protective antigen of
Bacillus anthracis."
RL Gene 69:287-300(1988).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=sterne;
RC MEDLINE=994545463; PubMed=10515943;
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,
Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,
Martinez Y., Rieke D., Svensson R., Jackson P.J.,
RT "Sequence and organization of pXOI, the large Bacillus anthracis

RT plasmid harboring the anthrax toxin genes.";
RN J. Bacteriol. 181:6509-6515(1999).
RN [3]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=Ames / isolate Florida / A2012;
RX MEDLINE=22061436; PubMed=12004073; DOI=10.1126/science.1071837;
RA Read T.D., Salzberg S.L., Pop M., Shumway M., Unayam L., Jiang L.,
Holtapple E., Busch J.D., Smith K.L., Schupp J.M., Solomon D.,
Keim P., Fraser C.M.,
RT "Comparative genome sequencing for discovery of novel polymorphisms in
Bacillus anthracis."
RL Science 296:2028-2033(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Ames ancestor;
RA Ravel J., Rasbo D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
Fraser C.M.,
RT "Bacillus anthracis comparative genomics."
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Membrane-bound (potential).
CC CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC
DR EMBL, W22589; AAA22636.1; ALT_INIT; Genomic DNA.
DR EMBL, AF065404; AAD32415.1; ALT_INIT; Genomic DNA.
DR EMBL, AB011190; AAM26108.1; -; Genomic DNA.
DR EMBL, AB017336; AAT28903.2; -; Genomic DNA.
DR PIR, G59104; G59104.
DR PIR, I39933; I39933.
DR HSSP, P13423; IACC.
DR GenomeReviews; AB017336 GR; GBAA_pXOI_0163.
DR TIGR, GBAA_pXOI_0163; -.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxinB; 1.
KW Complete proteome; Hypothetical protein; Membrane; Plasmid;
KW Transmembrane.
KW CHAIN 1 225
FT FT
FT FT
FT TRANSMEM 181 203
FT CONFLICT 114 114 T -> A (in Ref. 1).
FT CONFLICT 205 225 KSCNCLITVYEVSQLNMSVY -> NHVIVYLSM (in Ref. 1).
SQ SEQUENCE 225 AA; 25402 MW; 2E121B854295F9C8 CRC64;
Query Match 7.9%; Score 296.5; DB 1; Length 225;
Best Local Similarity 37.2%; Pred. No. 9.7e-07;
Matches 68; Conservative 38; Mismatches 54; Indels 23; Gaps 6;
QY 566 LAELNATNITVYLDKIKLANAKNIIILDKRPHYDRNNINAVGADSVYKAREVINSSTE 625
Db 1 MESLGINNIYNNLDRKIKNAKNIIIVRP-YHYDNGNIVGVDSYLNKAYQIILWSSD 59
QY 626 GLIANTDKDIRKILSGYIEIDTE-----GLKEVINDRYMNLNISLRDQ 672
Db 60 GVSINLDEVDNQALSGYMLQIKKPSNHLTNPSPTVITLAKDGSVGLYVLS-----DG 113
QY 673 KNPIDFKKYNKDLPLYISNPNKYVNVYATKEN-TIINSENGDSTNGKILIFSKG 731
Db 114 TGFIDFNKFDENWRSLV-DPGDDVYVAYKDFNAVTRDENGNIA-NKLKNTLVLSGKI 171
QY 732 YEI 734
Db 172 KEI 174

RESULT 15
G6FNZ8_CANGA PRELIMINARY; PRT; 1305 AA.
ID G6FNZ8_CANGA PRELIMINARY; PRT; 1305 AA.
AC G6FNZ8_19-UTL-2004, integrated into UniProtKB/TrEMBL.

19-JUL-2004, sequence version 1.
DT 21-FEB-2006, entry version 20.
DE Similar to sp|P1753 Saccharomyces cerevisiae YNL250w RAD50 DNA repair
protein.
OS OrderedLocNames=CAGL007788g;
OC Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
CX NCBI_TaxID=5478;
NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 2001 / CBS 138;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
Lafontaine I., de Montigny J., Marc C., Neveglisse C., Talla E.,
Goffard N., Frangeul L., Aigle M., Anthonard V., Babour A., Barde V.,
Barry S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
Bolsenne A., Boyer J., Catolico L., Confanieri F., de Daruvar A.,
Despons L., Fabre B., Fairhead C., Ferry-Dumazet H., Groppi A.,
Hantreay F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
Kerrest A., Koszul R., Lemaire M., Leaur I., Ma L., Muller H.,
Nicaud J.-M., Nikolski M., Ozias S., Ozier-Kalogeropoulos O.,
Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
Swennen D., Tekata F., Wesolowski-Louvel M., Westhof B., Wirth B.,
Zenou-Meyer M., Zivanovic Y., Bojotin-Fukuhara M., Thierry A.,
Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissbach J.,
Wincker P., Souciet J.-L.,
RA "Genome evolution in yeasts."
RT Nature 430:35-44(2004).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
EMBL, CR380956; CAG6097.1; -, Genomic DNA.
DR GO: GO:0005694; C:chromosome; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0030870; C:Mei1 complex; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0004518; F:nuclease activity; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR GO: GO:0051276; P:chromosome organization and biogenesis; IEA.
DR GO: GO:0006281; P:DNA repair; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003439; ABC transp_like.
DR InterPro: IPR004584; Rad50.
DR InterPro: IPR007517; Rad50_Zn_hook.
DR InterPro: IPR003395; SMC_N.
DR InterPro: IPR013134; Zn_hook_Rad50.
DR Pfam: PF04423; Rad50_zn_hook; 1.
DR Pfam: PF02463; SMC_N_1.
DR TIGRFAMs: TIGR00606; rad50; 1.
DR Complete proteome.
SQ SEQUENCE 1305 AA; 15066 MW; 7489860D317CER80 CRC64;
Query Match 6.2%; Score 232.5; DB 2; Length 1305;
Best Local Similarity 20.5%; Pred. No. 0.012; Mismatches 264; Indels 279; Gaps 39;
Matches 175; Conservative 135;
1 EVKQENRLIN-----ESESQGLGLGYFSDINFPAPVVTSTTGLSPSSR 49
225 KYKQINBKINDVQGVDDVEQELSESINGSLDLPFSNQNQFQSVLRIFETIANSTISAATQ 284
50 LENTDS-----EN--QYFQSAIWSGFIVYKKSDEYTFATPSADNHTVMVDDOE 95
285 IERLNSIEKLDMPKYDLENILKNFSESIEQKELDVKRLESKLAS-----DRQK 333
96 VINKASNGKIRLKEGRLYQIKIYQYR-----ENPTKGLDFK 133
334 LENCKQKQSRRLMLQOGBELPAKQDQNTLTKSPQDTIRQIKDHEKLDLNAEPPQRRVDHT 393
134 LYWTDSQNKKEVYISSDNLQLPELKQKSSNSRKRRSTSGAPTVPRDRNDGIDPDSLEVEG-- 191

394 L-----KKKADMESSVNEAISKLEETKGNLKEK-----SDSIRNSINLEBQK 435
192 ---YTVYVK-----NKRTFLSPWISNIHEKGLITKYSPEKMSSTASDPYSPPEKVTG 241
436 LSYVESDILKQATLDDQKQISSIPADIE-----QKAKKAWSKLSKELQEPK--- 487
242 RIDKRVSEARHPLVAAYPIVHVDMENIISKNDQ-----TQNTDSQRTTISKNTSTS 296
488 --ENMIT-----DINQIKSNBELSKSELQAQILETDLQNYSKSKSLTS 529
297 KDANTVGVSIASIQNGFTGNTTSAGFSNNSSTVAIDHSLAG-RTAETMTGLNTA 355
530 TKLSVL--IENNDEKQNEI--NKITFTALINDTMIXEKGILDVKQLSFEKTYIALQNKIA 586
356 DTALNINIRVYNGTAPIYVNVLPPTSILVGNKOT-----LATIKAKENQLSQ-----IL 405
587 ---TKLAFNE-----ISKQTELELEYELNTCKGKVNLIQHKKEKIL 625
406 APNNYPS---KNLAPIALNAODDFSGT-----PTMANYQFLBLEKTKQRLDTDQ 454
626 NINSVLPEDCSIEDYDVLLETSEVSYKTALENLQKHQTTLEFNR-----KALEVAINDC 680
455 VYGNIAIYVNFENGKRVVDTSNMSEVLPOIQETTRARITFNGKDLNLYERRIAAVNPSPDL 514
681 CY--LCRSFENTEFR-----SKILKELEKKT-----DKFEE-----SL 713
515 ETTKPDMLKEALKIARGFENPNGNLQYQKDIETEPDFNPDQTSQ----- 560
714 KTTLEDE--KEYL-----NNLRLAEKDI-----YNLSISSESALQDRISQLSKE 757
561 --NIKQQLAEIATNTIYTVLDKIKLNAKMNIIRDREPHYDRNNIYAGADE----- 609
758 LDNKKSEVAEAN-----TTIEKTK-----EKRDHCD-NYIKPMKDIDYLQKEIY 801
610 --SVYKAREVINSSTEG-----LLANIDKQIRKLSGIYIEIETEGKAVINDRY 660
802 NHSEKSKSLNDVIRSSAENGDSVTMBQLODNQKSTRELIMKIRSEIBSLQORENITSSKN 861
661 DMLNSSLRQ-----DGKTFIDPKKYNDK---PLYISNPYKVVAVYTKEN 705
862 NTL--INQIRQANDVAETIEKQFDMKVTIQEQIANDKHAIBELTNSKESIKLHINDLSKV 920
706 TIINPENGDTST 718
921 EMLKENSEKIST 933
Db

Search completed: August 11, 2006, 20:26:06
Job time : 315 secs